

Cuk- Hobart- 21- P2267

THE BRITISH
JOURNAL OF
EDUCATIONAL
PSYCHOLOGY

FEBRUARY 1980

VOL. 50, PART 1



THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by

Professor DENNIS CHILD

Assistant Editors:

Professor S. N. BENNETT

Dr. G. BROWN

Professor K. WEDELL

Sub-Editor:

Dr. S. CUNNINGHAM

Advised by the Editorial Committee:

Dr. H. BELOFF

Professor R. A. KENNEDY

Professor P. LEVY

Professor J. D. NISBET

Professor E. A. PEEL

Mr. P. PRINCEP

Professor B. SHACKEL

Dr. D. SHARPLES

Professor M. SUTHERLAND

Vol. 50, Part 1

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £12.00 a year post free, payable in advance, or £5.00 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication should be sent to Professor Dennis Child, School of Education, St. Thomas' Street, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All *Books for Review* for the *British Journal of Educational Psychology* should be sent to: Dr. G. Brown, Department of Educational Research, The University of Lancaster, Bailrigg, Lancaster.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by

SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1980. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY R. & R. CLARK LTD., EDINBURGH

Vol. 50

Pt. 1-3

1980

RIAL

92

21

All those who are concerned with the production of the *Journal* would like to join me in thanking Professor Noel Entwistle for devoting much time and effort as editor over the last five years. The task, like all others on the editorial side, is voluntary and requires vigilance in guarding the high standards achieved by the *Journal*. This he has done admirably. During his tenure he has introduced from time to time symposia on current issues and fostered an international flavour both in the range of acceptable papers from abroad and in the *Journal's* circulation. I shall certainly want to continue to encourage good quality papers from overseas whenever these are relevant to educational psychology.

Professor Neville Bennett and Dr. Geoff Brown have agreed to continue as assistant editors. In addition, Professor Klaus Wedell has agreed to join us as a third assistant. This, I believe, adds strength to our team by having someone with interests in the field of children with special needs. Dr. Shirley Cunningham has kindly accepted my invitation to continue as sub-editor.

The unabated flow of papers for consideration and the extensive circulation of the *Journal* are a testimony to the significant place enjoyed by the publication. Changes in publishing policy are not at present envisaged and will only take place in response to the development of the subject and the reaction of readers. As with previous editors, my particular interests in psychology will not in any way influence the selection of papers. We rely on many dependable referees representing the widest range of interests and we are grateful for their continued help. The quality of the *Journal* depends upon the choice of papers we receive and I want to encourage authors to send material to us. The competition for acceptance is still fierce, but that makes acceptance all the more rewarding.

I was intrigued to discover recently that the *Journal* actually started its life here at Newcastle-upon-Tyne. It was launched by one of my predecessors in the chair of education, Professor Mark Wright, who was the first Professor of Education in England. It appeared in 1908 as the *Training College Record*, but the title changed several times. In 1911 it became the *Journal of Experimental Pedagogy* and in 1923 *Forum of Education*. Its present title was adopted in 1931 under the editorship of Professor Valentine. I hope the next cycle of seventy years will be as successful as the first.

DENNIS CHILD

P22.69

NON-DIRECTIVE AND INTRUSIVE TEACHING STYLES OF MIDDLE- AND WORKING-CLASS ENGLISH MOTHERS

By JEAN S. PHINNEY

(*California State University, Los Angeles*)

AND NORMA DEITCH FESHBACH

(*University of California, Los Angeles*)

SUMMARY. Sixty-three mother-child dyads, selected from middle- and working-class English backgrounds, were observed in semi-structured teaching activities. The five activities were selected for the 2½-3½ year-old child sample. Observers scored precoded maternal and child behaviours. An analysis of covariance, with mother's age and child's age as covariates, indicated that middle-class mothers used significantly fewer negative statements and imperatives, and significantly more questions. They were also less intrusive. Within-class correlations showed two maternal styles in the working class, a negative intrusive style and a positive, non-directive style.

INTRODUCTION

In recent years, a considerable body of research has accumulated demonstrating the extensive relationship of social class to many aspects of the child's development. Membership in a given social class appears to entail a high degree of probability that a child will be exposed to a certain cluster of experiences different from those of a child from another class (Hess *et al.*, 1968; Deutsch, 1973). It further appears that these different experiences may be related to resulting differences in abilities, attitudes, and behaviours of the developing child (Hess, 1969). However, in order to get beyond the broad influence of social class and understand the psychological processes that may mediate different behavioural outcomes, recent research has focused on particular components of the child's experience. As Deutsch (1973, p. 241) points out in a recent review, the dissection out of the SES matrix of certain variables, in order to examine their effects separately, places research on social class differences on a more valid scientific basis.

One important component of the child's early environment is the manner in which the mother directs and teaches the child in the myriad of day-to-day interactions. Social class differences in maternal teaching styles have been well documented in recent research (Brophy, 1970; Streissguth and Bee, 1972). Specifically, lower-class mothers tend to be more controlling and directive than middle-class mothers (Zegib and Forehand, 1975); they tend to use less positive reinforcement and more negative or punitive statements than do middle-class mothers (Bee *et al.*, 1969; Brophy, 1970; Feshbach, 1973a). These findings have been shown to be consistent across a number of different cultures (Feshbach, 1973b). Greenglass (1972) found similar differences between immigrant and second generation Italian women living in Canada, suggesting that the second generation mothers had adopted the middle-class norms of their new country. The teaching behaviours of young children themselves show similar class differences (Feshbach, 1973a). In addition, lower-class and less educated mothers tend to be more physically intrusive when teaching their child (Bee *et al.*, 1969; Minton *et al.*, 1971). Robinson and Arnold (1977) found that middle-class mothers differ from working-class mothers in the way that they answer young children's questions; for example, they tend to relate answers to past experiences and extend the answers beyond the immediate question. On the basis of these data, Robinson suggests that "middle-class mothers can be considered more effective teachers."

However, while SES may be correlated with specific maternal behaviours, maternal styles may vary widely within a given class and have an effect independent of class. For example, a pattern of intrusiveness and negative reinforcement was found

to characterise middle-class mothers whose children were problem readers (Bercovici and Feshbach, 1973). Deutsch (1973) notes that for many variables within-class differences exceed differences between classes. Thus, the present study was aimed at examining maternal teaching styles within as well as across social class. In order to explore these issues in a cultural setting in which social class is generally independent of ethnic group, the study was carried out with a sample drawn from English middle- and working-class populations.

An important question in studying maternal behaviour is the extent to which behaviour observed in any one situation is characteristic of the many day-to-day interactions between mother and child. The generality of teaching behaviours across situations is likely to have a greater impact on the child. Gordon (1976) found mothers to be highly consistent across two occasions in structured play episodes with the same child, and also across two separate play episodes with two different siblings. The present study focuses on mother-child interactions in a variety of play situations in the home. The specific activities which were observed varied widely from a highly structured activity to a relatively unstructured one. Thus the mother has an opportunity to exhibit a number of behaviours that might be presumed typical of her everyday interactions with the child. A central question of the study was the extent to which these behaviours were consistent across the different activities.

Another relevant question pertains to the interrelationship among maternal behaviours, that is, the extent to which they occur in clusters and constitute distinct patterns of interaction. On the basis of previous research, it was assumed that negative and punitive behaviours would be related to the mother's directiveness (Bercovici and Feshbach, 1973). The present study was aimed at examining the extent to which these relationships are coherent patterns of behaviour across different situations.

METHOD

Sample

The sample observed consisted of 63 mother-child dyads, recruited through playgroups and waiting lists for admission to nursery schools in middle- and working-class neighbourhoods in Bristol, England.

Designation of social class followed a procedure previously used in assessing social class with an English population (Cook-Gumperz, 1973), adapted for the present purpose of identifying English middle- and working-class samples. Four criteria were used: mother's education, father's education, father's occupation, and residential area. The working-class sample consisted of 26 mother-child dyads who met three of the four following criteria: (1) mother left school at the minimum age of 15 or 16; (2) father left school at the minimum age of 15 or 16; (3) father was an unskilled or semi-skilled labourer; and (4) residential area was clearly working-class, generally public housing. The middle-class sample ($N=19$) met three of the four following criteria: (1) mother completed college or university; (2) father completed college or university; (3) father held a professional job; and (4) residential area was clearly middle-class, generally well-maintained private homes. An intermediate group ($N=18$) consisted of dyads with mixed or intermediate criteria, e.g., non-professional white-collar worker; some higher education; mixed residential neighbourhood.

The child's age, attendance at playgroup, and number of siblings did not differ significantly among class samples.

Procedure

Each mother-child pair was observed in their own home at a time arranged for the mother's convenience. The experimenter explained briefly that she had a number of activities and games for the mother to carry out with her child. The mother was

instructed to behave as she ordinarily would when alone with her child. The experimenter presented each activity to the mother with minimal instructions. In virtually all cases, mother and child played together on the living-room floor, becoming quickly involved in the activities, which were as follows. (1) Toy village: the child was given a collection of small houses, trees, farm animals and fences. The mother was asked to have the child make a toy village resembling one made by the experimenter. (2) Sorting: the mother was asked to have the child put away the pieces from the toy village by sorting them into four bags. (3) Puzzle: a nine-piece wooden jigsaw puzzle was placed in front of the child on the floor with the pieces removed; no instructions were given. (4) Game: a simple colour-matching game was explained to the mother who then played the game with the child. (5) Cuisenaire rods: an assorted collection of Cuisenaire rods was presented to the child. The child at first played freely with the rods. The mother was then given, in turn, three simple coloured patterns drawn on cardboard and was asked to have the child make the patterns with the rods.

The exact time spent on each activity and whether it was completed or not was recorded. In addition, the following precoded behaviours were scored at 30-second intervals, with a check made for the occurrence of each behaviour during the interval:

Mother's verbal behaviour. (1) Positive statements or praise; (2) negative, punitive or critical statements; (3) questions; (4) imperatives; and, (5) neutral or information-giving statements.

Mother's non-verbal behaviour. (1) Pointing or touching toys; (2) moving pieces, handing them to child, or placing them in position; (3) physically guiding or restraining the child.

Child's behaviour. Child not attending to activity, e.g., child persistently looks away from activity or becomes involved in some other activity.

Observations were made by two trained observers, one of them a graduate student attending the University of Bristol. Reliability between observers was assessed twice during data collection. During the initial assessment, reliability averaged 0.85 with a range of 0.75 to 0.94. At the second assessment, reliabilities increased to an average of 0.89 ranging from 0.81 to 0.98. Subsequent to the experimental procedures, a brief interview was carried out, to obtain demographic and social class data.

RESULTS

Scores for each behaviour were obtained by totalling the number of checks made for each behaviour on each activity and dividing by the number of time units (30-second intervals) spent on the activity yielding a ratio score adjusted for time differences. Since a preliminary analysis revealed no sex differences on any of the observed behaviours, the data for boys and girls were combined.

In order to examine the consistency of each mother's behaviours across activities, correlations were calculated between all activities for each of the observed maternal behaviours. Of the 80 correlations, all were positive except for two that were near zero; 52 were significant (see Table 1). The highest consistencies were obtained for negative statements, imperatives, moving pieces, and physical guidance.

In view of the consistency of maternal behaviours, scores for each behaviour were totalled across activities. The means and standard deviations for all the observed behaviours are shown in Table 2 for the middle-class, working-class, and intermediate groups. Since age of the mother was significantly related to social class (mean age of middle-class mothers: 31.9 years; of working-class mothers: 27.6 years; of intermediate groups: 28.8 years), and since significant correlations were obtained between a child's age and a number of maternal behaviours (as discussed below), analyses of covariance, using mother's age and child's age as covariates, were carried out to test for class differences in observed maternal behaviours.

TABLE 1
INTERCORRELATIONS (PEARSON r) AMONG TASKS FOR MATERNAL BEHAVIOURS

	Village with				Puzzle with			Sort with		Game with Rods
	Puzzle	Sort	Game	Rods	Sort	Game	Rods	Game	Rods	
Positive Non-Directive										
Positive Statements	0.50***	0.15	0.20	0.13	0.24	0.17	0.37**	0.11	0.18	0.34**
Neutral Statements	0.51***	0.18	0.18	0.32**	0.23	0.33**	0.14	0.09	0.02	-0.01
Questions	0.36**	0.38**	0.31*	0.58***	0.37*	0.24	0.28*	0.31*	0.20	0.20
Negative Intrusive										
Negative Statements	0.21	0.50***	0.53***	0.49***	0.32**	0.36**	0.42***	0.52***	0.49***	0.49***
Imperative	0.44***	0.52***	0.29*	0.59***	0.34*	0.07	0.54***	0.28*	0.66***	0.42***
Pointing	0.15	0.21	0.32**	0.31*	0.12	0.04	-0.03	0.11	0.19	0.29*
Moving	0.41***	0.28*	0.42***	0.53***	0.27*	0.39**	0.62***	0.15	0.26*	0.48***
Physical Guidance	0.10	0.49***	0.35**	0.59***	0.27*	0.42***	0.32**	0.47***	0.50***	0.64***

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$

TABLE 2

MEANS AND STANDARD DEVIATIONS OF OBSERVED BEHAVIOURS FOR THE WORKING, INTERMEDIATE, AND MIDDLE-CLASS GROUPS

	Working Class N=26		Intermediate Class N=18		Middle Class N=19		
Maternal Behaviours	Mean	SD	Mean	SD	Mean	SD	P
Positive Non-directive							
Positive Statements	54.0	16.7	56.1	17.1	60.7	14.3	NS
Neutral Statements	66.0	16.5	70.5	12.1	76.9	8.4	NS
Questions	43.0	19.0	50.2	19.7	70.9	10.2	0.001
Negative Intrusive							
Negative Statements	59.6	20.1	36.7	18.5	33.9	13.8	0.001
Imperatives	80.9	13.2	59.7	18.0	58.1	19.6	0.001
Pointing	64.5	17.9	53.3	11.9	51.2	14.1	NS(0.07)
Moving	53.8	21.3	39.6	19.5	30.1	10.9	0.001
Physical Guidance	16.6	18.5	6.7	11.6	4.1	6.8	0.05
Total Verbal	303.6	47.9	273.1	44.7	300.6	30.6	NS
Total Non-verbal	134.9	46.6	99.6	33.8	85.4	18.0	0.001
Child's Performance							
Child Not Attending	11.0	17.3	6.8	9.4	8.1	9.0	NS
No. of Tasks Completed	4.9	1.6	5.2	1.6	5.0	1.6	NS
Time to Complete Tasks	7.7	1.9	7.1	2.3	7.7	2.0	NS

The results reflect significant class differences in the use of negative statements ($P < 0.001$), imperatives ($P < 0.01$), questions ($P < 0.001$), moving pieces ($P < 0.01$) and total non-verbal behaviour ($P < 0.001$). Middle-class mothers used fewer negative statements, fewer imperatives, asked more questions, and were less apt to move pieces. Also, middle-class mothers tended to use less physical guidance, although this class difference fell short of statistical significance ($P < 0.07$). There was no significant difference in total amount of verbal behaviour, but working-class mothers exhibited more non-verbal behaviour. For all maternal behaviour, except use of imperatives,

the working-class sample showed greater variability. An $F_{\max}$ test for the assumption of homogeneity of variance was carried out for all observed maternal behaviours (Kirk, 1968). The assumption of homogeneity was supported for positive and negative statements, imperatives, and pointing, and rejected for questions, neutral statements, moving pieces, and physical guidance. According to Kirk, the F distribution is robust with respect to violation of the assumption of homogeneity of variance. In terms of the child's performance, it is noteworthy that there were no significant class differences in time taken to complete the activities, number of activities completed, or the child's attention to the activities.

Pearson product-moment correlations were carried out to determine the inter-relationships among the maternal behaviours. Because of the differences between classes, correlations among observed behaviours were calculated for the middle- and working-class groups separately (Table 3). The resulting positive and negative correlations suggest a clustering of the behaviours into two patterns. A pattern of negative intrusive behaviour is shown by the positive intercorrelations among negative statements, imperatives, pointing, moving pieces, and physical guidance and low or negative correlations with positive, non-directive behaviours. The positive, non-directive behaviours, which include positive and neutral statements and questions, show significant positive intercorrelations in the working-class but not the middle-class. It is particularly interesting that the working-class sample demonstrated both the positive non-directive and negative intrusive patterns, while in the middle-class sample only the negative intrusive pattern was evident. These within-class correlations suggest clearly that maternal styles have a coherence within as well as between social classes.

TABLE 3

PEARSON PRODUCT-MOMENT CORRELATIONS OF MOTHER'S BEHAVIOURS FOR WORKING-CLASS AND MIDDLE-CLASS GROUPS

Maternal Behaviour	Positive non-directive			Negative intrusive			
	Positive	Neutral	Questions	Negative	Imperat.	Point	Move
<i>A. Working-class Mothers</i>							
Positive Non-directive							
Positive							
Neutral	0.44**						
Question	0.23	0.54***					
Negative Intrusive							
Negative	0.18	0.07	-0.03				
Imperatives	-0.07	-0.13	-0.33	0.35			
Point	0.13	0.18	0.03	0.60***	0.26		
Move	-0.10	-0.15	-0.21	0.39	0.47***	0.49**	
Guide	-0.30	-0.35	-0.52	0.34	0.46**	0.35	0.57**
<i>B. Middle-class Mothers</i>							
Positive Non-directive							
Positive							
Neutral	-0.25						
Question	-0.31	0.07					
Negative Intrusive							
Negative	-0.04	0.12	0.02				
Imperatives	0.21	-0.37	-0.43	0.52**			
Point	0.02	-0.30	0.21	0.03	0.23		
Move	0.14	-0.31	-0.39	0.13	0.35	-0.26	
Guide	0.36	-0.01	-0.20	0.34	0.45†	0.20	0.23

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ † $P < 0.055$

The relationship of the mother's behaviour to several demographic variables and to the child's performance was examined by means of Pearson product-moment correlations among these factors and the observed behaviours (Table 4). The child's age appeared to have some influence on maternal behaviours, with the mothers of younger children showing more negative intrusive behaviour. Another interesting demographic relationship was that mothers of children who attended playgroups or attended more often tended to be less intrusive and more positive than mothers whose children did not attend playgroups or attended less frequently. As previously noted, neither age of the child nor attendance at playgroup differed significantly among the three socio-economic groups.

TABLE 4
CORRELATIONS OF DEMOGRAPHIC AND PERFORMANCE VARIABLES WITH MOTHER'S TEACHING STYLE

Maternal Behaviours	Demographic				Child's Performance		
	Mother's Age	Mother's Education	Child's Age	Child to Playgroup	Child not Attending	No. Tasks Completed	Time to Complete
Positive Non-directive							
Positive	0.31*	0.07	0.14	0.34***	-0.51***	0.33**	-0.28*
Neutral	0.19	0.23	0.15	0.43***	-0.28*	0.08	0.08
Questions	0.09	0.51***	0.12	0.10	-0.15	-0.15	0.08
Negative Intrusive							
Negative	-0.25*	-0.52***	-0.29*	-0.24	-0.06	-0.16	0.08
Imperatives	-0.22	-0.48***	-0.22	-0.35**	0.15	-0.10	0.09
Pointing	-0.33**	-0.31*	-0.22	-0.32*	-0.14	0.07	-0.11
Moving	-0.01	-0.53***	-0.25	-0.23	0.21	-0.01	-0.13
Physical Guidance	-0.13	-0.37**	-0.38**	-0.41***	0.30**	-0.07	0.06

* $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$

The child's performance showed a certain relationship to the mother's behaviour, but less than could be anticipated. The child who was inattentive received fewer positive statements and more physical guidance and had fewer neutral statements directed at him/her. However, inattentiveness was not significantly related to negative statements, imperatives, questions, pointing, or moving pieces. The child's completion of activities was related only to positive statements among the maternal behaviours. Thus, in general, the mother's behaviours, particularly the negative intrusive behaviours, were less a response to what the child did than they were a reflection of the mother's own consistent style.

DISCUSSION

The social class differences found in this study with English mothers of three-year-olds are consistent with findings from previous studies carried out with older children. We find in the verbal sphere that middle-class mothers use fewer negatives, fewer imperatives, and ask more questions than working-class mothers. Among the non-verbal behaviours, the clearly intrusive behaviours—moving pieces and physically guiding the child—occur more often in the English working class.

It is possible that the mothers are responding to differences in the children's competence, since differences in cognitive functioning among children from different socio-economic backgrounds emerge between eighteen months and three years of age (Caldwell, 1971). However, while the child's cognitive competence was not independently assessed, his/her attention to the activity, the number of activities com-

pleted, and the time to complete each activity showed no significant differences between social groups. Furthermore, among the maternal behaviours, only positive statements were consistently related to the child's performance. The maternal behaviours were generally independent of the child's performance, a result consistent with that of Minton *et al.* (1971). Thus social class differences in maternal styles cannot be readily attributed to differences in the child's performance.

Not surprisingly, the education of the mother appears to be an important factor mediating class differences, with the better educated mothers showing fewer of the negative intrusive behaviours, in accord with previous research (e.g., Minton *et al.*, 1971). Furthermore, the child's attendance at playgroup, while not significantly different among classes, was related positively to the non-directive style and negatively to the intrusive style. The mother who sends her child to playgroup may already be more concerned with the child and better informed about child care, or it may be that through participation in the playgroup, the mother develops a more positive way of interacting with the child.

Perhaps the most intriguing of our findings is the reflected pattern of two distinct maternal teaching styles that show up within and across classes, i.e., a positive, non-directive style that includes positive and neutral statements and questions, and a negative intrusive style consisting of negative statements, imperatives, moving pieces, and physical guidance. These patterns appear in the significant intercorrelations among the various behaviours within classes and also in the consistency of a given behaviour across a number of different activities. Despite the variation among these activities in their structure and cognitive demands, maternal reactions displayed a marked commonality. It seems likely that the behaviours manifested by these mothers in the semi-structured experimental situations may be prototypic of the patterns of behaviour exhibited outside the experimental situation.

Of the two patterns, the negative intrusive pattern is relatively more independent of the child's performance. The negative intrusive mother seems to exhibit similar behaviour regardless of the child's performance and may thus be less sensitive in responding to him/her. In contrast, positive statements are clearly related to the child's performance; the mother who uses a positive style may thus be responding more contingently to the child.

Evidence is accumulating that components of what we have called an intrusive style interfere with the child's cognitive development. Hess and his colleagues (cited in Hess, 1969) noted that school achievement was positively related to the mother's avoidance of imperative strategies. Nelson (1973) reports that intrusiveness of the mother interferes with language acquisition and concept formation. Recently Donovan (1975) showed that a mother's verbal commands without explanation related to poor performance by a child on a number of representational tasks. It thus seems likely that a mother's pervasive use of a negative intrusive style can be detrimental to a child's mental growth.

Moreover, the negative intrusive style has a punitive quality that has been related to problems in personality and social development. For example, Hess (1969) notes that mothers who used imperative strategies had sons who tended to get low conduct grades. Thus a pattern of negative intrusive behaviours, compared to a more positive style, may mediate both poor school performance and conduct problems among children whose mothers use this style consistently, regardless of the social class to which they belong.

ACKNOWLEDGMENTS.—This research was supported in part by a grant from the UCLA Committee on International Comparative Studies. Thanks are extended to Naomi Roberts for her assistance in collecting data.

REFERENCES

- BEE, H. L., EGEREN, L. F., STREISSGUTH, A. P., NYMAN, B. A., and LECKIE, M. S. (1969). Social class differences in maternal teaching strategies and speech patterns. *Dev. Psychol.*, 1, 726-734.
- BERCOVICI, A., and FESHBACH, N. D. (1973). Teaching styles of mothers of successful and problem readers. Paper presented at the meeting of the American Educational Research Association, New Orleans, February, 1973.
- BROPHY, J. (1970). Mothers as teachers of their own preschool children: the influence of socio-economic status and task structure on teaching specificity. *Child Dev.*, 41, 79-94.
- CALDWELL, B. (1971). The rationale for early intervention. In COHEN, S. (Ed.), *Child Development: A Study of Growth Processes*. Itasca, Ill: Peacock.
- COOK-GUMPERZ, J. (1973). *Social Control and Socialization: A Study of Class Differences in the Language of Maternal Control*. London: Routledge and Kegan Paul.
- DEUTSCH, C. (1973). Social class and child development. In CALDWELL, B. E. and RICCIUTI, H. N. (Eds.), *Review of Child Development Research*. Vol. 3. Chicago: University of Chicago Press.
- DONOVAN, A. (1975). The relationship between maternal behavior and representation in children. Paper presented at the meeting of the Society for Research in Child Development, Denver, April 1975.
- FESHBACH, N. D. (1973a). Teaching styles in four-year-olds and their mothers. In ROSENBLITH, J. R. and ALLINSMITH, W. (Eds.), *The Causes of Behaviours: Readings in Child Development and Educational Psychology* (3rd ed.). Boston: Allyn and Bacon.
- FESHBACH, N. D. (1973b). Cross-cultural studies of teaching styles in four-year-olds and their mothers. In PICK, A. (Ed.), *Minnesota Symposia on Child Psychology*. Minneapolis: University of Minnesota Press.
- GORDON, D. (1976). Consistency of mother's behavior in repeated play interactions with their children. *J. genet. Psychol.*, 129, 337-338.
- GREENGLASS, E. (1972). A comparison of maternal communication style between immigrant Italian and second-generation Italian women living in Canada. *J. Cross-Cultural Psychol.*, 3, 185-189.
- HESS, R. D. (1969). Family characteristics and school achievement in young children. In GROTEBERG, E. (Ed.), *Critical Issues in Research Related to Disadvantaged Children*. Princeton, N.J.: Educational Testing Service.
- HESS, R. D., SHIPMAN, V. C., BROPHY, J., and BEAR, R. (1968). *The Cognitive Environments of Urban Preschool Children*. Chicago: University of Chicago Press.
- KIRK, R. E. (1968). *Experimental Design: Procedures for the behavioral sciences*. Belmont, Calif.: Brooks/Cole.
- MINTON, C., KAGAN, J. M., and LEVINE, J. (1971). Maternal control and obedience in the two-year-old. *Child Dev.*, 42, 1873-1894.
- NELSON, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 38 (1-2), Serial No. 149.
- ROBINSON, W. P., and ARNOLD, J. (1977). The question-answer exchange between mothers and young children. *European J. soc. Psychol.*, 7, 151-164.
- STREISSGUTH, A. P., and BEE, H. L. (1972). Mother-child interaction and cognitive development in children. In HARTUP, W. (Ed.), *The Young Child*, Vol. 2. Washington, D.C.: National Association for the Education of Young Children.
- TULKIN, S., and KAGAN, J. (1972). Mother-child interaction in the first year of life. *Child Dev.*, 43, 31-41.
- ZEGIOB, L. E., and FOREHAND, R. (1975). Maternal interactive behavior as a function of race, socio-economic status, and sex of the child. *Child Dev.*, 46, 564-568.

(Manuscript received 5th July, 1979)

SOME FACTORS CONTRIBUTING TO INTELLIGENCE AT AGE OF SCHOOL ENTRY

BY P. A. SILVA AND J. BRADSHAW

(*University of Otago, Dunedin*)

SUMMARY. This study reports the contribution of maternal general ability, some developmental milestones, early motor development, language development, and experiences to intelligence as measured at age five. The results confirm that there is a relatively low level of prediction of intelligence as the multiple correlation of all the independent variables is only 0.61. The implications of the low level of the variance of intelligence explained is discussed.

INTRODUCTION

DESPITE the fact that students of human development have long sought to gain a better understanding of factors contributing to the variation in intelligence (e.g., Hunt, 1961; White, 1971), there is still remarkably little known about the precise influence of parental, experiential and developmental factors on child intelligence. The problem was recently restated by Clarke and Clarke (1978) who noted that there is, as yet, no proper model of human development. The complexity of the problem was underscored by their general statement that . . . "human development is an ongoing process of interaction and transaction between genetic/constitutional factors on one hand and macro and micro social factors on the other" (p. 3).

The study reported in this paper represents a further examination of some factors contributing to child intelligence and is an extension of a study reported by Silva and Fergusson (1976). In the earlier paper reporting results from an earlier sample, the influence of socio-economic status, some maternal characteristics, and child experience on child intelligence at age four years was described and discussed in the context of the 'nature-nurture' controversy. Altogether, five independent variables were studied and these resulted in a multiple correlation of 0.54 with child intelligence. Because of the relatively low multiple correlation obtained, the authors emphasised that a more adequate understanding of factors influencing intelligence requires the consideration of more factors than were included in the earlier study. It was decided, therefore, to examine whether the addition of further child experience measures and some measures of early motor and language developmental status would result in a significantly higher multiple correlation with child intelligence.

As in the earlier study, the dependent variable selected was child intelligence as measured by the Stanford-Binet Intelligence Scale (Terman and Merrill, 1960). This measure was used because of the wealth of reliability and validation data available and especially because it is known to be a good predictor of scholastic success (e.g., Anastasi, 1961).

The independent variables selected for study were the two most 'important' ones from the Silva and Fergusson (1976) study (child experiences and maternal general mental ability), some additional experience measures, and five early motor and language developmental status measures. The total of nine independent variables was selected on the basis of evidence from previous studies (Silva, 1976; Silva and Fergusson, 1976; Silva, 1978a) which suggested that their combined effect could possibly result in a larger multiple correlation with child intelligence.

METHODS

Sample

The sample consisted of 797 of the 1,037 children who are members of the Dunedin Multidisciplinary Child Development Study sample. This has been described elsewhere (Silva, 1976; Silva, 1978a). Only those for whom all the measures were

complete were included, as the method of analysis required the availability of all the results for all children.

Measures

The measures were taken when the children were within a month of their third and fifth birthdays. These are described in detail in Silva (1976). The measures are briefly described below:

Child intelligence. This was based on the Stanford-Binet Intelligence Scale (Terman and Merrill, 1960) administered when the children were within a month of their fifth birthdays.

Maternal general mental ability. This was assessed by the SRA Verbal Test (Thurstone and Thurstone, 1973), the results of which are described by Silva (1978b). It was administered to the mothers when their children were within a month of their third birthdays.

Child experiences. Two measures of child experiences were used. The Experiences Scale (Silva, 1976; Silva and Fergusson, 1976) was administered when the children were within a month of their third birthdays and again when the children were within a month of their fifth birthdays. The scale consists of a check-list of 30 experiences the children may have had in the two-year period prior to assessment. The other experiences measure was the number of half days each five-year-old child had attended a pre-school educational facility.

Early developmental status. Five measures were taken:

- (1) Age in months when children first walked.
- (2) Age in months when children first talked in sentences.

All the mothers were asked to bring their Plunket Books which record when the milestones were achieved. Most of the losses from the sample were due to the mothers being uncertain when the milestones were achieved. The Plunket Society provides a free infant and pre-school nursing service to oversee child health and development of most New Zealand children and the Plunket Books record pertinent developmental information such as the age at which children first walk and talk.

- (3) Motor development assessed at age three with the Bayley (1969) Motor Scales.
- (4) Verbal comprehension.
- (5) Verbal expression measured at age three by the Reynell Developmental Language Scales (Reynell, 1969; McKerracher *et al.*, 1977).

ANALYSIS AND RESULTS

Product moment correlations followed by stepwise regression analysis (Kerlinger and Pedhauster, 1974) were used to analyse the data.

The means and standard deviations for the measures are set out in Table 1 and the correlations among the variables are set out in Table 2.

TABLE 1
MEANS AND STANDARD DEVIATIONS FOR THE MEASURES (N = 797)

Measure	Mean	SD
1. Stanford-Binet IQ (5 yrs)	107.2*	15.96
2. Maternal general ability	40.5	14.87
3. Age child walked (mths)	13.5	2.70
4. Age child talked (sentences—mths)	24.2	5.58
5. Motor score (3 yrs)	20.6	4.16
6. Verbal comprehension (3 yrs)	35.5	8.45
7. Verbal expression (3 yrs)	36.4	8.09
8. Child experiences (3 yrs)	18.4	3.44
9. Child experiences (5 yrs)	21.2	3.67
10. Educational experiences (5 yrs)	175.3	82.12

* The IQ is based on the Thorndike (1972) norms.

TABLE 2
CORRELATIONS AMONG THE VARIABLES (N=797)

Measures	1	2	3	4	5	6	7	8	9
1. Stanford-Binet IQ (5 yrs)	—								
2. Maternal general ability	0.31	—							
3. Age child walked	-0.12	0.00	—						
4. Age child talked (sentences)	-0.18	-0.06	0.10	—					
5. Motor score (3 yrs)	0.22	0.06	-0.47	-0.08	—				
6. Verbal comprehension (3 yrs)	0.54	0.22	-0.28	-0.30	0.35	—			
7. Verbal expression (3 yrs)	0.48	0.17	-0.22	-0.33	0.28	0.60	—		
8. Child experiences (3 yrs)	0.23	0.18	-0.13	-0.17	0.16	0.22	0.22	—	
9. Child experiences (5 yrs)	0.30	0.24	-0.09	-0.14	0.15	0.27	0.21	0.61	—
10. Educ. experiences (5 yrs)	0.11	0.03	-0.09	0.00	0.04	0.11	0.07	0.11	0.11

All the measures correlated with child intelligence at a level that exceeded statistical significance. The two measures of three-year language ability correlated most highly (0.48 and 0.54), followed by maternal general mental ability (0.31), and child experiences at age five years (0.30).

The stepwise regression analysis on child intelligence at age five is summarised in Table 3.

Only the first four variables in the analysis resulted in F ratios that were significant beyond the 0.01 level. The multiple correlation for these four variables with intelligence was 0.61 and the percentage of the variance of the intelligence measure explained was 37.48.

TABLE 3
SUMMARY OF STEPWISE REGRESSION ANALYSIS ON CHILD INTELLIGENCE (N = 797)

Step No.	Variable	Multiple R	% of Variance Explained	Increase in % of Variance Explained	F Ratio to Enter or Remove
1.	Verbal comprehension (3 yrs)	0.54	29.09	—	258.08
2.	Maternal general ability	0.57	32.84	3.76	38.81
3.	Verbal expression (3 yrs)	0.60	36.21	3.37	36.60
4.	Child experiences (5 yrs)	0.61	37.48	1.26	14.00
5.	Age child talked	0.61	37.64	0.17	1.84
6.	Age child walked	0.61	37.80	0.16	1.77
7.	Educational experiences (5 yrs)	0.62	37.94	0.13	1.50
8.	Motor score (3 yrs)	0.62	38.05	0.12	1.30
9.	Child experiences (3 yrs)	0.62	38.10	0.05	1.00

DISCUSSION

The independent variables which correlated most highly with the child intelligence measure at age five were the three-year language measures (0.54 and 0.48). It was of interest to note that the correlation obtained between the three-year verbal comprehension measure and child intelligence at age five was of the same magnitude as the multiple correlations obtained for all the independent variables at age four and reported in the Silva and Fergusson paper (1976). This finding emphasises the relative importance of pre-school language development and suggests that one direct observation of developmental language status in the pre-school years is as good a predictor of intelligence at this age as is a study of the combined effect of socioeconomic status, maternal general mental ability, maternal training in child development, maternal authoritarian attitudes, and child experiences. However, neither is especially powerful.

While the three-year measures of language development were designated 'developmental status measures', it should be acknowledged that these measures may well be considered as measuring general mental ability as well, a view that has been substantiated by a factor analytic study including the Reynell Scales and a number of other measures (Silva *et al.*, in press).

As in the Silva and Fergusson (1976) study, maternal general ability and the five-year experiences measure correlated with five-year intelligence at a moderate level (0.31, 0.30), each explaining slightly less than 10 per cent of the variance of intelligence. The three-year measure of child experiences resulted in a lower univariate correlation with five-year intelligence (0.23) than the five-year experience measure (0.30), although these two correlations are not significantly different. It was of interest to find that these two experience measures were themselves moderately highly correlated (0.61) despite the fact that there was a two-year interval between administrations of the scales.

In contrast to the moderate correlations between the experiences scale measures and child intelligence was the relatively low (but statistically significant) correlation between the pre-school education experience measures and child intelligence (0.11). In this study, less than 1 per cent of the variance of intelligence was explained by attendance at pre-school and it contributed nothing extra in the regression analysis. This suggests that more time at a pre-school does not contribute substantially to child intelligence. It is also of interest that maternal intelligence was not related to length of time at pre-school.

The early developmental status measures of age of walking and talking resulted in only small (but statistically significant) correlations (-0.12 , -0.18) with child intelligence, as did the direct measure of child motor ability at age three years (0.22). These results suggest that quite early motor and language developmental status are related to later intelligence but the relationship is very weak.

Clarke (1978) stated that "We live in a world where everything seems to have some relation with everything else; it is extremely difficult to avoid nature's favourite correlation co-efficient, about 0.35. It is because of this that the employment of many variables additively or in weighted combination normally fails to improve prediction very greatly. After the combination of two or three variables, the law of diminishing returns nearly always operates" (p. 252).

This study gives strong support to Clarke's statement; the variables that were used to predict intelligence tended to be intercorrelated (some very near to 0.35) and addition of more than a few variables did not greatly improve the prediction. The variable that was most highly correlated with intelligence was three-year verbal comprehension (0.54 or 29 per cent of the variance explained); addition of all the other variables only raised this to 0.61, or 38 per cent of the variance explained. In other words, addition of further variables only increased the proportion of the variance explained by 9 per cent. Thus it is concluded that addition of a wider range of variables does *not* result in a greatly increased explanation of the variance of intelligence.

While the low level of explanation of child intelligence resulting from this study may, on the surface, seem disappointing, it does serve as a timely reminder of the *low* levels of explanation and prediction of school entry intelligence that are currently possible. The findings serve to emphasise that students of human development will need to continue their search for a better understanding of factors that contribute to child intelligence. Further parameters will need to be studied and alternative measures will need to be developed before substantial advances in knowledge of the problem can be achieved.

More needs to be learned about the role of parental characteristics, child experiences and early developmental characteristics and how they relate to the development of intelligence which is, itself, poorly understood. Clarke (1978) does not believe

that poor prediction can be blamed on inadequacies in our methods of measurement. Rather, he believes they are due to the nature of human development which he describes as being characterised by inconstancy rather than constancy, particularly in the pre-school years.

In addition to discovering about the role of factors influencing child intelligence is the need to learn more about the *interactions* between these factors and the child and how the child himself influences his environment and experiences.

The results of this study should serve as a reminder that predictions of school entry intelligence on the basis of knowledge about their mother's intelligence, about the child's experiences or about their early developmental characteristics are not likely to be reliable. This study has shown that the predictive validity of these factors, either singly or additively, is very low when studied in a large group of pre-school children. Moreover, the period over which prediction could be made was relatively short. All the factors included in this study explained only 38 per cent of the variance of child intelligence at the age of school entry; 62 per cent of the variance remained unexplained. This suggests that clinicians should take a conservative and open-ended view of the role of factors that might influence intelligence, reserving judgment and avoiding predictions. If predictions are poor at the group level, predictions about individuals are likely to be even less reliable.

The reason for the relatively low level of prediction in the present study probably lies partly in the age at which intelligence was assessed. It is widely held that development during the early years proceeds at a relatively fast rate (e.g., Bloom, 1964) and fluctuates fairly rapidly during this period (Clarke, 1978). Further changes in a large proportion are to be expected throughout childhood and it will be of interest to try to establish some of the parental, experiential, and other correlates.

ACKNOWLEDGMENTS.—The Dunedin Multidisciplinary Child Development Study is supported by the Medical Research Council of New Zealand, the New Zealand Department of Education, the Department of Health, and involves several departments of the University of Otago. Much of the data is gathered by volunteer workers from the Dunedin community. The authors are indebted to the many people whose contribution makes this ongoing study possible.

REFERENCES

- ANASTASI, A. (1961). *Psychological Testing*. New York: Macmillan.
- BAYLEY, N. (1969). *The Bayley Scales of Infant Development*. New York: The Psychological Corporation.
- BLOOM, B. S. (1964). *Stability and Change in Human Characteristics*. New York: John Wiley.
- CLARKE, A. D. B. (1978). Predicting human development: problems, evidence, implications. *Bull. Brit. Psychol. Soc.*, 31, 249.
- CLARKE, A. M., and CLARKE, A. D. B. (1978). The limited later effects of early experience. A paper presented to the NZ/OECD Conference on Early Childhood Care and Education, Massey University, February, 1978.
- HUNT, J. McV. (1961). *Intelligence and Experience*. New York: Ronald Press.
- KERLINGER, F. M., and PEDHAUSER, E. J. (1974). *Multiple Regression in Behavioural Research*. New York: Holt, Rinehart and Winston.
- McKERRACHER, D. W., SAKLOFSKE, D. H., and SILVA, P. A. (1977). An evaluation and cross cultural comparison of the Reynell Developmental Language Scales. *Australian Reading Education Journal*, Spring, 1977; 14, 2.
- REYNELL, J. (1969). *Reynell Developmental Language Scales*. London: National Foundation for Educational Research.
- SILVA, P. A. (1976). *A Thousand Dunedin Three-Year-Olds*. Unpublished research report presented to the Medical Research Council of New Zealand. Available from the University of Otago Medical Library.
- SILVA, P. A. (1978a). Some biological and environmental factors contributing to early child development. A brief description of the Queen Mary Multidisciplinary Child Development Study. In O'Rourke, B., and Clough, J. (Eds.), *Early Childhood in NZ*. Auckland: Heinemann Educational.

- SILVA, P. A. (1978b). SRA Verbal Test scores from 1011 women. *NZ Psychologist*, 7, 47.
- SILVA, P. A., BRADSHAW, J., and SPEARS, G. F. A study of the concurrent and predictive validity of the Reynell Developmental Language Scales. Windsor: NFER. In press.
- SILVA, P. A., and FERGUSON, D. M. (1976). Socio-economic status, parental characteristics, child experience and intelligence in pre-school children: a Path Analytic Model. *NZ J. educ. Studies*, 11, 180-188.
- TERMAN, L. M., and MERRILL, M. R. (1960). *Stanford-Binet Intelligence Scale*. Boston: Houghton Mifflin.
- THORNDIKE, R. L. (1972). *Stanford-Binet Intelligence Scale: 1972 Norms Tables*. Boston: Houghton Mifflin.
- THURSTONE, T. G., and THURSTONE, L. L. (1973). *The SRA Verbal Form*. Chicago: Science Research Associates.
- WHITE, B. L. (1971). *Human Infants: Experience and Psychological Development*. New Jersey: Prentice-Hall.

(Manuscript received 11th June, 1979)



THE ITPA VISUAL SEQUENTIAL MEMORY TASK: AN ALTERNATIVE INTERPRETATION AND THE IMPLICATIONS FOR GOOD AND POOR READERS

By CAROLYN HICKS

(Applied Psychology Department, University of Aston in Birmingham)

SUMMARY. Four experiments were carried out to examine the different recall strategies employed in a diagnostic test of visual sequential memory (VSM). Briefly the results were as follows: Experiment 1 indicated that competent readers tend to use a verbal labelling strategy in the recall of visual stimuli, rather than visual memory *per se*. Experiment 2 suggested that the retention of the visual stimuli could be improved by the adoption of a verbal labelling strategy. Experiment 3 suggested that when verbal labelling was suppressed the performance of competent readers on the VSM task deteriorated to a level similar to that of poor readers. Experiment 4 suggested that if retarded readers were instructed to use a verbal labelling strategy their retention of visual symbols improved significantly. Taken together, the principal implication from these results is that good and poor readers may differ not with respect to visual memory but their differential ability to employ a verbal labelling strategy in the retention of visual stimuli.

INTRODUCTION

THE plethora of research surrounding the psychological aspects of reading has featured, to a large extent, the identification of the pre-requisite sub-skills necessary for the acquisition of literacy. One methodological approach within this field has involved the comparison of good and poor readers on all these sub-skills (e.g., Spring, 1971; Zangwill and Blakemore, 1972; Guthrie, 1973; Vellutino *et al.*, 1975). While any combination of failure in these sub-skills would probably have a detrimental effect on reading performance, it is conceivable that some skills may be more important than others to the reading process. To isolate these, tasks aimed at testing the abilities of good and poor readers on each reading sub-skill have been developed.

Of particular relevance here is the work done on the role of the visual processing function in reading. Considerable evidence has been amassed to support the hypothesis that good and poor readers differ with regard to the amount of information they can extract from a visual display (Lyle, 1968; Lyle and Goyen, 1968; Vernon, 1971; Guthrie and Goldberg, 1972). Furthermore, it has been demonstrated that this information can be coded in different ways. Coltheart (1972) has suggested that, following transfer of information from iconic memory to a different form of memory function, recall may be based upon a stored label of the item or upon a stored visual code, the former being of greater capacity and negligible decay. Information stored by means of a verbal label is more readily available for rehearsal and reproduction than that which has been processed by a visual code (den Heyer and Barrett, 1971; Scarborough, 1972).

If visual information can be stored in either of these two ways then it is conceivable that studies involved in the evaluation of visual memory may have been assessing retention of verbal labels, rather than visual memory *per se*, and this may especially be the case where stimulus items are highly codable verbally (Lyle, 1968; Lyle and Goyen, 1968; Guthrie and Goldberg, 1972). Therefore, any differences between good and poor readers on such tasks may be attributable to the subject's ability to employ verbal labelling strategies to encode the visual information, rather than to actual differences in short term visual memory. Therefore, to test the possibility that differences in visual memory are related to reading attainment, it would seem necessary to devise a task where verbalisation is completely eliminated. While Bakker (1967) may

have gone some way towards developing such a task, the majority of those tests currently in clinical diagnostic use do not appear to avoid the potentially confounding effects of verbal labelling.

One test frequently used in the psychometric assessment of reading failure is the Illinois Test of Psycholinguistic Abilities (ITPA). This is a diagnostic test divided into 12 sub-tests, each of which is designed to isolate specific sub-skill failure underlying psycho-linguistic learning disabilities. Of particular interest here is the visual sequential memory sub-test. (The visual sequential memory task involves the presentation of a card on which is printed a number of arbitrary visual symbols. The testee is required to look at the card for five seconds, and upon its removal to replicate the sequence using individual plastic symbols.) A number of studies have noted the diagnostic and predictive ability of this task in relation to reading disabilities (Kass, 1966; Hirshoren, 1969; Macione, 1969) and conclusions have consequently been drawn as to the role of visual sequential memory (VSM) in the reading process (de Hirsch *et al.*, 1966; Vernon, 1971; Kirk and Kirk, 1971). While it is possible that visual sequential memory *is* a vital function in the reading process, it is conceivable, alternatively, that what is being measured in this task is not visual memory, but the retention of verbal labels assigned to the visual symbols.

The present study was aimed at establishing:

- (a) what the ITPA VSM task measures;
- (b) what skills are required for adequate test performance;
- (c) how far good and poor readers differ in these skills.

The hypothesis examined in Experiment 1 relates to the alternative explanation of the ITPA VSM results, that is, the task may not be a measure of visual short term memory but rather of the subject's ability to label the symbols, and that subjects who label the stimuli are more likely to recall more of the visual display.

EXPERIMENT 1

Subjects

Twenty children participated in the experiment of which ten were male and ten were female. The mean age was 9 : 4 years (range 8 : 11-9 : 10 years) and the mean reading age as measured by the Schonell Word Recognition Test was 9 : 6 years. All the children came within the average intelligence range, as measured by the Wechsler Intelligence Scale for Children.

Materials

The visual sequential memory task from the ITPA was used; this consists of a stopwatch, a display booklet and a number of individual plastic chips, each with an arbitrary visual symbol.

Procedure

The Schonell Graded Word Recognition Test was given prior to the experiment, in order to eliminate subjects who did not achieve a pre-established minimum level of competence (less than three months below CA). The administration of the visual sequential memory task was in accordance with the instructions in the ITPA manual. The task was given to each subject individually; the 17 chips were randomly arranged in front of the subject and the subject was given one minute in which to look at them (pre-exposure period). The relevant sequence card in the display booklet was then exposed for five seconds after the experimental instructions had been completed. On removal of the display card, the subject was required to replicate the sequence using the corresponding chips. If the subject completed the first trial, no second trial was administered; if the subject failed the first trial, a second attempt was permitted after re-exposure of the sequence card.

To score the test, two marks were awarded for a correct response on the first trial and one mark for a correct response on the second trial.

The sampling procedure recommended for subjects over four years old was adopted. Using this approach, only pre-selected items were administered, until the subject failed on the first trial of any item. The second trial was then given, but the sampling procedure was discontinued. Each display not previously administered was then given in reverse order, until three consecutively listed items were passed on the first trial. The experimenter then returned to the point in the sampling procedure where the first failure occurred. Consecutive test items were presented thereafter, until two items were failed on both trials.

In the present study, on completion of the task, each subject was asked to report the nature of his recall strategy in order that:

- (1) It could be established whether or not, under standard assessment conditions, strategies of recall other than that intended (i.e., visual memory) were used.
- (2) The performance of subjects who used a verbal labelling code could be compared with the performance of subjects who used a visual code.

RESULTS

Of the 20 subjects used, 13 reported using a verbal labelling strategy, while the remaining seven reported recall based on the configuration of the stimulus items. Those subjects who reported using a verbal labelling strategy said that the symbols reminded them of common objects, and named the symbols accordingly. The children reported rehearsing the order of the names during the presentation of the sequence. Subjects who reported using a visual recall strategy said that they tried to picture the shape in their mind. Further exploration of this by the experimenter indicated that no label had been attached and that recall was based on retention of the outline shape of the symbols.

These results suggest that under normal testing conditions, the ITPA VSM task does not necessarily measure short term visual memory. Using a *t*-test for unrelated samples, the difference in scores between the two groups is significant at the 0.005 level ($t = 3.57$, $df = 18$). This indicates that subjects using a verbal labelling strategy recall more of the visual symbols than subjects who use a visual store (see Table 1).

The subjects who used a visual recall strategy may have adopted this approach simply to comply with the experimental instructions and the apparent nature of the task. It is conceivable that the natural recall strategy of these competent readers also involved naming the stimuli, and that the perceived role demands of the experiment led to a temporary strategy change.

While the experiment has not clarified the essential differences between good and poor readers on the ITPA VSM task, it has demonstrated that the task may be invalid as a test of visual short term memory, since the possibility of verbally encoding the stimulus items cannot be completely ruled out. Furthermore, it has indicated that this verbal strategy is more efficient in the retention of visual information.

If age equivalent scores on this task depend on the use of a verbal coding strategy, then it may be that a critical factor which differentiates high and low scorers on this test is the ability to transfer visual information into a more durable verbal store. If low scoring subjects are provided with this strategy, then it should be possible to improve their performance on the task. Furthermore, it was predicted that subjects who had previously used a verbal labelling strategy would not show a significant improvement in test scores.

EXPERIMENT 2

Subjects and materials

The subjects and materials from Experiment 1 were used in the present study.

Procedure

On completion of Experiment 1, the subjects were required to re-perform the ITPA VSM task, but with the following modifications:

- (1) During the pre-exposure period, the subjects were required to assign verbal labels to each of the visual symbols. For instance, the symbol  was usually labelled 'star'. If after this period, the subject had failed to name any of the symbols, a label was provided by the experimenter. These labels used the most common responses supplied by the verbal recall group of Experiment 1, on the assumption that they would be more meaningful to the subjects than labels imposed by the experimenter. For example, the symbol  was named 'Catherine Wheel'.
- (2) The experimental instructions were modified so that the subject was specifically requested to recall the verbal, rather than visual, sequence.

With these exceptions, the procedure was identical to that of Experiment 1.

RESULTS

A 2×2 factorial analysis of variance for unequal samples yielded both significant main effects and a significant interaction between subject group and the labelling manipulation. This indicates that the two subject groups responded differentially to the experimental treatment. Using a *t*-test for related samples ($t = 1.289$, $df = 12$, NS) the results of the Name Code group for Experiments 1 and 2 indicate that the scores of the group did not improve significantly under the experimental instructions to label the stimulus items.

TABLE 1

A COMPARISON OF THE PERFORMANCE OF THE NAME CODE GROUP AND THE VISUAL CODE GROUP UNDER STANDARD ASSESSMENT PROCEDURES (EXPERIMENT 1) AND THE VERBAL LABELLING CONDITIONS (EXPERIMENT 2)

	Exp. 1	Exp. 2	$\bar{X}$
Name Code Group (N = 13)	22.69	23.23	22.96
Visual Code Group (N = 7)	19.71	22.71	21.21
$\bar{X}$	21.2	22.97	—

TABLE 2

SUMMARY TABLE OF 2×2 FACTORIAL ANALYSIS OF VARIANCE
(A = SUBJECT GROUP; B = EXPERIMENTAL TREATMENT)

Source	DF	SS	MS	F	P
A	1	27.782	27.782	11.916	<0.01
B	1	28.484	28.484	12.217	<0.01
A $\times$ B	1	13.781	13.781	5.911	<0.05
Within cell	36	83.934	2.331	—	—

However, when the results of the Visual Code group for Experiments 1 and 2 are compared using a related *t*-test ($t = 6.874$, $df = 6$, $P < 0.001$), it can be seen that these subjects performed significantly better under the experimental instructions to label the stimulus items.

A comparison of the results using an independent *t*-test for the Verbal Code and Visual Code groups in Experiment 2 indicates that there was no significant difference in scores when both groups labelled the stimulus items ($t = 0.921$, $df = 18$, NS).

If the results from the previous experiments could be shown to apply to poor readers, then they have important implications for the identification of critical sub-skills important for literacy attainment. While it is highly probable that visual memory does have an important function in the reading process, it is, nonetheless, difficult to evaluate its relative capacity in good and poor readers in view of the fact that many of the available diagnostic tests of visual memory do not consistently measure this factor. However, if it could be demonstrated that STVM is approximately the same regardless of reading ability, then any differences on the ITPA VSM task may be attributable to the operation of other sub-skills, namely verbal codability. More specifically, the hypotheses underlying the next experiment suggest that:

- (a) poor readers are more likely to use visual recall strategies on the ITPA VSM task than are good readers;
- (b) if sub-vocalisation of labelling responses of the visual stimuli could be eliminated by the introduction of an incompatible activity, the performance of subjects who had previously used a verbal code would deteriorate since they would be compelled to use the alternative visual memory store.

EXPERIMENT 3

Subjects

The subjects consisted of a further group of 12 children, six boys and six girls, mean age 9 : 7 years and mean reading age on the Schonell Graded Word Recognition Test of 9 years 10 months. In addition, a group of 12 dyslexics, ten boys and two girls, were used, mean age 9 : 4 years and mean reading age on the Schonell Graded Word Recognition Test of 7 : 8 years. Both groups came within the average intelligence range on the WISC.

Apparatus

As before.

Procedure

Stage 1

The first stage of the experiment was identical to the procedure used in Experiment 1; the ITPA visual sequential memory task was administered to each subject individually and on completion of the task the subject was asked to report the nature of his recall strategy. However, the experimental instructions were modified slightly such that the subjects were asked to recall the stimuli 'in any way you can' in an attempt to eliminate from the experiment the possible confounding effect of compliance with the perceived demands of the task. On the basis of the subjects' report of the recall strategy used, the subjects were divided into Name Code Group and Visual Code Group; these results formed a baseline measure with which to compare subsequent performance.

Stage 2

The task was re-administered with a further modification: the subject was asked to say the word 'the' as quickly as possible throughout the experimental procedure. It

was hypothesised that a concurrent competing verbal response would suppress the tendency to name the visual symbols, thereby eliminating the use of a verbal memory store. This procedure has been used extensively by Murray (1967, 1968) and Levy (1971).

It was predicted that:

- (1) the dyslexic group would be more likely to report using a visual recall strategy during stage 1 of the experiment;
- (2) those subjects who reported using a visual memory store should show no significant decrease in scores during the second stage of the task when sub-vocalisation was suppressed;
- (3) if the labelling response is fully suppressed then those subjects who reported using a verbal memory store should perform at a level equivalent to that of the Visual Code Group, i.e., their scores should deteriorate during the second stage of the task.

RESULTS

Of the 12 competent readers, only one subject reported using a visual recall strategy in stage 1, as compared with 9 of the 12 dyslexics. This suggests that poor readers may be more likely to use inefficient visual retention strategies than are competent readers.

The results of the experiment are presented in Tables 3 and 4. A 2×2 factorial analysis of variance for unequal samples yielded both significant main effects and significant interaction between subject group and verbal suppression. An independent

TABLE 3

A COMPARISON OF THE PERFORMANCE OF THE NAME CODE GROUP AND THE VISUAL CODE GROUP UNDER STANDARD ASSESSMENT PROCEDURES (EXPERIMENT 3, STAGE 1) AND VERBAL SUPPRESSION CONDITIONS (EXPERIMENT 3, STAGE 2)

	Exp. 3 Stage 1	Exp. 3 Stage 2	$\bar{X}$
Name Code Group (N = 14)	22.86	19.43	21.15
Visual Code Group (N = 10)	19.6	19.1	19.35
$\bar{X}$	21.23	19.27	—

TABLE 4

SUMMARY TABLE OF 2×2 FACTORIAL ANALYSIS OF VARIANCE
(A = SUBJECT GROUP; B = EXPERIMENTAL TREATMENT)

Source	DF	SS	MS	F	P
A	1	37.5	37.5	13.259	<0.01
B	1	45.015	45.015	15.916	<0.01
A $\times$ B	1	25.014	25.014	8.844	<0.01
Within cell	44	124.441	2.828	—	—

t-test revealed a significant difference in performance between the Name Code Group and the Visual Code Group ($t = 4.471$, $df = 22$, $P < 0.001$). The above results under standard assessment procedure, which form a baseline of measurement for the experiment, also show that:

P2269

- (1) the earlier findings are corroborated, in that subjects who store the stimulus items in a verbal rather than visual store, retain more visual information;
- (2) competent readers are more likely to use the more efficient verbal coding strategy than are poor readers.

Using a *t*-test for unrelated samples ($t = 0.500$, $df = 22$, NS) there was no significant difference between the performance of the Name Code Group and Visual Code Group under conditions of verbal suppression. These results suggest that when a competing verbal response is introduced to suppress stimulus labelling the scores of the Name Code Group deteriorate to a level equivalent to that of the Visual Code Group.

When the results of the Visual Code Group for stages 1 and 2 are compared using a *t*-test for related samples ($t = 1.246$, $df = 9$, NS), it can be seen that there was no significant difference in recall of visual stimuli for subjects who use visual coding strategies when a competing verbal response is introduced to suppress any sub-vocalisation of stimulus labelling.

When the results of the Name Code Group for Stages 1 and 2 are compared using a *t*-test for related samples ($t = 13.68$, $df = 13$, $P < 0.001$), it is shown that the performance of subjects who previously used a verbal storage system for visual stimulus items deteriorates when verbal labelling is suppressed.

The foregoing experiments have demonstrated that good and poor readers differ not with respect to their short term memory capacities but to their differential ability to transfer visual stimulus items into a verbal store. From this it might be predicted that if poor readers were given a general strategy for recalling visual information, namely by labelling the stimuli, then their recall on the ITPA VSM task should improve. This forms the hypothesis for the final experiment.

EXPERIMENT 4

Subjects

The subjects consisted of a further group of 12 dyslexic children, eight boys and four girls, mean age 9 years 3 months and mean age on the Schonell Graded Word Recognition Test of 7 : 8 years. A further group of 12 competent readers, seven boys and five girls, were used, mean age 8 years 11 months and mean reading age on the Schonell Graded Word Recognition Test of 9 : 2 years. All subjects came within the average intelligence range as measured by the WISC.

Apparatus

As before.

Procedure

The procedure was divided into two stages. Stage 1 involved administering the ITPA VSM task to each subject individually, and on completion of the task the subject was asked to report his recall strategy. Of the 12 competent readers, two reported using visual recall and so were excluded from the statistical analysis; all the dyslexic subjects reported using a visual recall strategy.

Stage 2 involved the re-administration of the task, but the subjects were asked to label the symbols during the pre-exposure period. If any subjects failed to do this, a label was supplied. The subjects were then asked to recall the sequence using the verbal labels.

RESULTS

The results are presented in Tables 5 and 6. A 2×2 factorial analysis of variance yielded both significant main effects and a significant interaction. A *t*-test for related

TABLE 5

A COMPARISON OF THE PERFORMANCE OF THE DYSLExIC GROUP AND THE NORMAL GROUP UNDER STANDARD ASSESSMENT PROCEDURES (EXPERIMENT 4, STAGE 1) AND VERBAL LABELLING CONDITIONS (EXPERIMENT 4, STAGE 2)

	Exp. 4 Stage 1	Exp. 4 Stage 2	$\bar{X}$
Dyslexics (N = 12)	13.84	17.09	15.47
Normals (N = 10)	19.8	20.2	20.0
$\bar{X}$	16.82	18.65	—

TABLE 6

SUMMARY TABLE OF 2 × 2 FACTORIAL ANALYSIS OF VARIANCE
(A = SUBJECT GROUP; B = EXPERIMENTAL TREATMENT)

Source	DF	SS	MS	F	P
A	1	225.016	225.016	48.974	<0.01
B	1	36.331	36.331	7.907	<0.01
A × B	1	22.154	22.154	4.822	<0.05
Within cell	40	183.783	4.595		

samples was used to compare the performance of the dyslexic subjects for Stages 1 and 2. The results are significant at the 0.001 level ($t = 9.892$, $df = 11$). This suggests that if poor readers are directed to use a more efficient storage strategy for retention of visual stimuli their results on a task of visual sequential memory improve significantly.

A comparison of the performance of good readers under standard assessment procedures and verbal labelling conditions showed no significant difference ($t = 0.768$; $df = 9$, NS).

When the scores of the good readers and the dyslexic readers from Stage 1 are compared, the results are highly significant ($t = 6.06$, $df = 20$, $P < 0.001$). When the scores of the good readers and the dyslexic readers from Stage 2 are compared, the results are significant ($t = 3.59$, $df = 20$, $P < 0.01$).

DISCUSSION

These four studies have indicated that the ITPA VSM task does not necessarily measure visual sequential memory, but rather verbal encoding ability. The results have indicated that good readers are more likely to label the stimulus items and to retain them in an acoustic memory store, whereas poor readers attempt recall based on visual memory. Because of the features which characterise auditory and visual short term memory (Neisser, 1967; Coltheart, 1972) subjects who transfer the visual information into a verbal memory store attain higher results on the ITPA VSM task. Furthermore, when competent readers who visually encode the information are directed to label stimuli, their recall capacity increases, confirming the comparative efficiency of this strategy. This instruction has no effect upon the performance of verbal encoders, indicating that they are already using an optimal recall strategy. The general pattern of results also applied to poor readers; however, while recall is improved

by using a labelling strategy, the amount of information retained is not equivalent to that retained by competent readers. A possible explanation for this relates to the limited short term memory capacity of dyslexics (Thomson and Wilsher, 1978). Since the ITPA VSM task tends to assess different skills across subjects, its reliability and validity as a diagnostic test of literacy failure must be questioned. Furthermore, on this basis, the norms of the test may be unrealistic, particularly in view of the fact that the verbal encoders achieved scores on the task which were commensurate with their age norms. It would therefore appear that the age norms for visual memory are based on verbal labelling (and hence, *verbal* short term memory) ability. Conversely, when subjects are prevented from labelling the stimulus items by a verbal suppression technique there is a deterioration in the performance of verbalisers but not in that of visualisers. This suggests that the acoustic encoding aspect can be eliminated from a visual memory task, using this method. Furthermore, the fact that the performance of visualisers did not deteriorate under conditions of verbal suppression indicates that this experimental treatment does not serve as a distractor.

In addition, this is compounded by the high codability of the stimulus items. In the experimental conditions designed to assess verbal coding most subjects had very little difficulty in labelling the symbols. To achieve a pure test of visual memory, then, it would appear to be necessary to devise stimulus items which do not resemble easily definable objects. However, with the methodological problems inherent in this, it may be preferable to further explore the possibilities of verbal suppression on existing visual memory tasks, especially in the light of the results obtained in the present study.

In spite of its shortcomings, the ITPA VSM task does appear generally (if inconsistently) to differentiate between good and poor readers. While this variation had been previously attributed to differences in visual short term memory capacity, a more plausible explanation (in terms of the present results) relates to the differential ability of good and poor readers to transfer visual information into a more durable verbal memory store. When verbal labelling is eliminated from the recall process, good and poor readers do not appear to differ significantly in their visual memory capacity. This has important ramifications for diagnosis and prescription based on VSM results. Assumptions about the relative capacity of STVM based on such results has led to the development of remedial programmes designed to overcome the inferred deficits.

However, since visual memory is characterised by quick decay and limited capacity it would seem to be of little value to train the failing reader in visual short term memory skills. Instead, it might be more profitable to provide general strategies for transferring information into a verbal memory store. This visual/auditory sensory integration has been demonstrated to be a critical determinant of reading failure (Birch and Belmont, 1965*a* and *b*; Ford, 1967; Reilly, 1971) and remediation of deficits in intermodal functioning should lead theoretically to greater achievement in literacy development than simply training visual memory. Research conducted to examine the efficacy of visual short term memory training on subsequent reading skills corroborates this (Bieger, 1978).

While further research both into specific sub-skill failure in reading and its subsequent remediation are necessary it would appear that perhaps for too long some aspects of literacy difficulties have been erroneously attributed to visual memory deficits.

ACKNOWLEDGEMENT.—The author would like to thank Dr D. C. Mitchell and Peter Jackson for their comments on an earlier draft of this paper.

REFERENCES

- BAKKER, D. J. (1967). Temporal order, meaningfulness and reading ability. *Percept. mot. Skills*, 24, 1027-1030.
- BIEGER, E. (1978). Effectiveness of visual training of letters and words on reading skills of non-readers. *J. educ. Res.*, 71, 157-161.
- BIRCH, H. G., and BELMONT, L. (1965a). Auditory-visual integration in brain-damaged and normal children. *Dev. Med. child Neurol.*, 7, 135-144.
- BIRCH, H. G. and BELMONT, L. (1965b). Auditory-visual integration, intelligence and reading ability in children. *Percept. mot. Skills*, 20, 295-305.
- COLTHEART, M. (1972). Visual information processing. In DODWELL, P. C. (Ed.), *New Horizons in Psychology II*. Harmondsworth: Penguin.
- DE HIRSCH, K., JANSKY, J. J. and LANGFORD, W. S. (1966). *Predicting Reading Failure: A Preliminary Study of Reading, Writing and Spelling Disabilities in Pre-School Children*. New York: Harper and Row.
- DEN HEYER, K., and BARRETT, B. (1971). Selective loss of visual and verbal information in short-term memory by means of visual and verbal interpolated tasks. *Psychon. Sci.*, 25, 100-102.
- FORD, M. P. (1967). Auditory-visual and tactual-visual integration in relation to reading ability. *Percept. mot. Skills*, 24, 831-841.
- GUTHRIE, J. T. (1973). Reading comprehension and syntactic responses in good and poor readers. *J. educ. Psychol.* 65, (3), 294-299.
- GUTHRIE, J. T. and GOLDBERG, H. K. (1972). Visual sequential memory in reading disability. *J. learning Disab.*, 5 (1), 41-46.
- HIRSHOREN, A. A. (1969). A comparison of the predictive validity of the revised Stanford-Binet Intelligence Scale and the Illinois Test of Psycholinguistic Abilities. *Except. Children*, 35, 517-521.
- KASS, C. E. (1966). Psycholinguistic disabilities of children with reading problems. *Except. Children*, 8, 533-539.
- KIRK, S. A., and KIRK, W. D. (1971). *Psycholinguistic Learning Disabilities: Diagnosis and Remediation*. Urbana, Ill.: University of Illinois Press.
- LEVY, B. A. (1971). The role of articulation in auditory and visual short term memory. *J. verb. Learn. verb. Behav.* 10, 123-132.
- LYLE, J. G. (1968). Performance of retarded readers on the Memory-for-Designs test. *Percept. mot. Skills*, 26, 851-854.
- LYLE, J. G. and GOYEN, J. (1968). Visual recognition, developmental lag and strephosymbolia in reading retardation. *J. abnorm. Psychol.*, 73, 25-29.
- MACIONE, J. R. (1969). Psychological correlates of reading disability as defined by the Illinois Test of Psycholinguistic Abilities. Doctoral Dissertation, University of South Dakota.
- MURRAY, D. J. (1967). The role of speech responses in short term memory. *Can. J. Psychol.*, 21, 263-276.
- MURRAY, D. J. (1968). Articulation and acoustic confusability in short term memory. *J. exper. Psychol.*, 78, 679-684.
- NEISSER, U. (1967). *Cognitive Psychology*. New York: Appleton-Century-Crofts.
- REILLY, D. M. (1971). Auditory-visual integration, sex and reading achievement. *J. educ. Psychol.*, 62, 482-486.
- SCARBOROUGH, D. L. (1972). Memory for brief visual displays: the role of implicit speech. *Cognit. Psychol.*, 3, 408-429.
- SPRING, C. (1971). Perceptual speed in poor readers. *J. educ. Psychol.*, 62, 492-500.
- THOMSON, M. E., and WILSHER, C. R. (1978). Some aspects of memory in dyslexics and controls. In GRUNEBERG, M. M., MORRIS, P. E., and SYKES, R. N. (Eds.), *Practical Aspects of Memory*. London: Academic Press.
- VELLUTINO, F. R., STEGER, J. A., DE SETTO, L., and PHILLIPS, F. (1975). Immediate and delayed recognition of visual stimuli in poor and normal readers. *J. exper. Child Psychol.*, 19, 223-232.
- VERNON, M. D. (1971). *Reading and its Difficulties: a Psychological Study*. Cambridge: Cambridge University Press.
- ZANGWILL, O. L. and BLAKEMORE, C. (1972). Dyslexia: reversal of eye-movements during reading. *Neuropsychol.*, 10, 371-373.

(Manuscript received 15th August 1979)

ADJUSTMENT FROM PRIMARY TO SECONDARY SCHOOL: A ONE YEAR FOLLOW-UP

By J. R. DOWLING*

(School Psychological Service, South Glamorgan)

SUMMARY. Four hundred and eighty-two children aged 10 and 11 years who were in their final year in primary school were followed up a year later to determine the extent to which personality, behavioural and educational variables as measured by tests, questionnaires and ratings, related to their social, emotional and educational adjustment in the secondary school. In general, correlations between the primary school variables and criterion variables were low. Behaviour in the primary school showed the highest correlation with adjustment after transfer whilst personality measures were unrelated to the criteria. Multiple regression analyses showed that the predictive efficiency of the primary variables in combination was also poor.

INTRODUCTION

ALTHOUGH educationalists would probably agree that the transfer from primary to secondary school is a significant and potentially difficult stage in a child's development, surprisingly little systematic research has been done into the changes of adjustment that may arise after transfer. The literature on the subject appears to be largely somewhat speculative (e.g., Blyth, 1965; Douglas, 1967; Plowden Committee, 1967; Halsall, 1975).

Probably the largest empirical investigation into the transfer is that of Nisbet and Entwistle (1969), but in using attainment after transfer as the main outcome measure the study throws little light on the effect of the transfer on social and emotional adjustment. Youngman's (1978) study into the different individual reactions to the transfer includes non-intellectual variables amongst the secondary measures. However, the omission of any measure of social and emotional behaviour reduces the value of the study in some respects.

Small scale studies by Ball (1975) and Hawthorn (1976) indicate that for some children the transfer appears to precipitate school adjustment problems not previously observed whilst for others the transfer seems to have a positive effect on adjustment. However, the findings of Cox (1978) suggest that perhaps the incidence of change of adjustment from top primary to first-year secondary is no greater than over any other 12-month period during the school years.

Although many researchers have found significant relationships between early antecedent variables and later childhood adjustment (e.g., Ullman, 1952; Gronlund and Holmlund, 1958; Pimm and McClure, 1967; Westman *et al.*, 1967; Davie *et al.*, 1972; Wedge and Prosser, 1973), relationships were not strong enough to carry any practical implications for educationalists. In these studies, however, children were followed up over developmental periods that were generally different from, and considerably longer than, that in the present investigation.

Teacher or peer ratings of behaviour have been found to be the most efficient predictors of future adjustment (Ullman, 1957; Bower, 1969; Liem *et al.*, 1969; Yellott *et al.*, 1969; Cowen *et al.*, 1973) but Pumfrey and Ward (1976) in a four-year follow-up from primary to secondary school showed that other psychometric variables i.e., intelligence and personality, combined with behaviour ratings in a multiple regression analysis did add significant increments to predictive efficiency.

The aim of the present study was to determine to what extent data obtained during children's final year in primary school, which is the normal traditional time when various kinds of assessment and prognosis are made, are related to their adjustment one year later after transfer to secondary school.

* Now with ILEA Schools' Psychological Service, 388/396 Oxford St., London W1.

METHOD

Sample

Five hundred and three children, aged between 10 years 6 months and 11 years 5 months who were in their final year in primary school, were selected. The children were from six junior schools (age range 7-11 years) which fed three co-educational comprehensive schools (age range 11-18 years). The schools were selected to try to obtain an approximately normal distribution of ability. However, there were no children ascertained as educationally subnormal in the sample.

Four hundred and eighty-two children (227 boys, 255 girls) were located a year later in the secondary school. Nineteen out of the 21 missing from the original sample had moved out of the area whilst the other two had unexpectedly transferred to secondary schools not included in the study.

Variables

(a) *Primary school variables.* The choice of the nine variables and the instruments used to measure them was influenced by practical considerations as well as previous research findings. It was considered whether use could be made of any existing assessment procedures and whether the experimental procedure could be repeated annually by primary school teachers, if this was suggested by the results. Thus the Moray House test 91 which was already administered in the schools was selected as a measure of intelligence in preference to other tests that could have provided a more general assessment of intellectual ability. No social variable, e.g., social class, was included because of the problems of reliably obtaining the information to determine this as part of an on-going, annual assessment carried out by teachers in the schools.

The variables selected were:

1. Moray House Test 91—verbal reasoning (University of Edinburgh, 1974).
2. National Foundation for Educational Research Test SRB—sentence reading comprehension (NFER, 1971).
3. Junior Eysenck Personality Inventory—(i) introversion-extraversion, (ii) stability-neuroticism, (iii) lie scale (Eysenck, 1965).
4. Child Scale B(2)—behaviour (Rutter, 1967).
5. Teachers' predictive ratings—4-point predictive scale of adjustment.
6. Attendance—percentage for final year in primary school.
7. Sex.

The predictive rating scale was designed by the author for the study:

- 1=I think that this child will make a satisfactory social and emotional adjustment in form 1 of the secondary school.
- 2=I think that this child may have some slight difficulty in making a satisfactory social and/or emotional adjustment in form 1 of the secondary school.
- 3=I think that this child will have considerable social and/or emotional problems and find it very difficult to cope in form 1 of the secondary school.
- 4=I think that this child will have such severe social and/or emotional problems that he/she will be unable to cope in form 1 of the secondary school.

The scale was piloted on a sample of 51 top primary children from a school not included in the study proper. From independent ratings of the sample by two teachers, a Spearman reliability coefficient of 0.70 ($P < 0.001$) was obtained.

(b) *Criterion variables.* Adjustment to secondary school was measured by four different indices, namely, behaviour, adjustment ratings, attitude to school and attendance. Behaviour was again measured by Child Scale B (2). The secondary adjustment rating scale was again a 4-point scale with the verbal descriptions closely paralleling those used for the primary scale. Attitude to school was measured by a questionnaire constructed by Fitt (1956). Attendance was measured as a percentage of possible attendances at the end of children's first year in the secondary school.

Procedure

The Moray House and SRB reading tests were administered by primary class teachers during one week in February prior to the children leaving the school in July. The Junior Eysenck Personality Inventory was also administered by the teachers during February. Class teachers completed a Child Scale B(2) for each child and rated each child during February and March.

The follow-up was done during April and May the following year, i.e., after the children had completed two terms in the secondary school. The assessment procedure in the secondary schools varied slightly according to the organisation of each school. However, in each case the child's form teacher gave the initial rating and completed Child Scale B(2). In one of the three schools the first-year tutor noted any discrepancies between her perception of the child and the form teacher's and they conferred before submitting the final assessment. In another of the secondary schools the form teachers conferred in the same way with the head of lower school.

The attitude questionnaires were given to all the children by the author personally. This was to try to obtain honest responses and the confidentiality of the responses was stressed.

RESULTS

Correlation

The Pearson correlations between the primary school and criterion variables are shown in Table 1.

TABLE 1
CORRELATION MATRIX (PEARSON) BETWEEN PRIMARY SCHOOL AND CRITERION VARIABLES

Criterion Variables		Predictor Variables								
		Sex	Teacher's Prediction	Verbal Reasoning Quotient	Reading Quotient	Extra-version	Neuroticism	Lie Score	Behaviour	Attendance
Rating	r	0.12	0.31	-0.20	-0.19	-0.02	-0.05	0.00	0.35	-0.18
	N	482	481	455	468	468	468	468	476	481
	sig. (2-tailed)	0.009	0.001	0.001	0.001	0.687	0.262	0.974	0.001	0.001
Behaviour	r	0.12	0.38	-0.24	-0.27	-0.07	-0.04	0.02	0.43	-0.14
	N	482	481	455	468	468	468	468	476	481
	sig. (2-tailed)	0.011	0.001	0.001	0.001	0.137	0.446	0.674	0.001	0.003
Attitude	r	0.18	0.17	-0.08	-0.10	0.07	0.08	-0.12	0.11	-0.06
	N	427	426	409	418	416	416	416	422	426
	sig. (2-tailed)	0.001	0.001	0.117	0.037	0.182	0.096	0.014	0.028	0.173
Attendance	r	0.08	-0.16	0.24	0.19	0.12	-0.08	-0.10	-0.21	0.69
	N	480	479	454	466	466	466	466	474	479
	sig. (2-tailed)	0.092	0.001	0.001	0.001	0.012	0.095	0.033	0.001	0.001

The teacher's predictive rating had highly significant correlations with all four criterion variables. Behaviour score (Scale B(2)) in the primary school had highly significant correlations with three of the criterion variables and a significant correlation with the fourth (attitude). In fact, overall, the behaviour score was the primary school variable most strongly related to adjustment in the secondary school in that three of the four coefficients were higher than those for teacher's prediction. Verbal reasoning quotient and reading quotient also had highly significant correlations with

three of the criterion variables and reading quotient was significantly correlated with the fourth (attitude). However, except for the correlation with attendance, the coefficients tended to be lower than those for behaviour and teacher's prediction.

The personality measures, extraversion and neuroticism, showed the weakest association with the criterion variables. Only one of the correlations was statistically significant—that between extraversion and attendance, i.e., introverts tend to have worse attendance records than extraverts.

Not surprisingly those variables that doubled as primary school and criterion measures showed relatively high correlations between primary and secondary scores. It can be seen that children's attendance is much more consistent across the transfer than behaviour (as measured by Scale B(2)).

Multiple regression

The stepwise regression of secondary teacher's adjustment rating with all the primary school variables is summarised in Table 2. With all the variables included only 17.1 per cent of the variance in secondary teachers' ratings is accounted for. On the third step (attendance) of the regression analysis the F value is 3.254 which indicates that the increase in multiple R does not reach the 5 per cent level of significance, (for $P < 0.05$, $F = 3.86$ for $df_1 = 1$ and $df_2 = 400$).

TABLE 2
SUMMARY OF THE STEPWISE REGRESSION OF SECONDARY TEACHER'S ADJUSTMENT
RATING WITH ALL THE PRIMARY SCHOOL VARIABLES

Primary School Variable	Multiple R	R^2	R^2 Change
1. Behaviour score	0.34791*	0.12104	0.12104
2. Teacher's predictive rating	0.38271*	0.14646	0.02542
3. Attendance	0.39149	0.15327	0.00680
4. Neuroticism score	0.39880	0.15904	0.00577
5. Sex	0.40340	0.16273	0.00369
6. Reading quotient	0.40825	0.16667	0.00394
7. Lie score	0.41215	0.16987	0.00320
8. Extraversion score	0.41395	0.17136	0.00149
9. Verbal reasoning quotient	0.41407	0.17146	0.00010

* $P < 0.05$

Table 3 summarises the stepwise regression of behaviour score (Scale B(2)) after transfer with all the predictor variables. It can be seen that with all the variables in the equation, approximately 24.5 per cent of the variance is accounted for. However, it is inefficient to proceed beyond step 3 as the marginal improvement of R^2 is so small.

TABLE 3
SUMMARY OF THE STEPWISE REGRESSION OF BEHAVIOUR SCORE (SCALE B(2))
AFTER TRANSFER WITH ALL THE PRIMARY SCHOOL VARIABLES

Primary School Variable	Multiple R	R^2	R^2 Change
1. Behaviour score	0.43228*	0.18687	0.18687
2. Teacher's predictive rating	0.46941*	0.22035	0.03348
3. Reading quotient	0.48141*	0.23175	0.01141
4. Sex	0.48771	0.23786	0.00611
5. Neuroticism score	0.49167	0.24174	0.00388
6. Verbal reasoning quotient	0.49297	0.24302	0.00128
7. Lie score	0.49423	0.24426	0.00124
8. Attendance	0.49508	0.24511	0.00084
9. Extraversion score	0.49534	0.24537	0.00026

* $P < 0.05$

The first four steps in the regression analysis of attitude to secondary school with all the primary school variables are shown in Table 4. The increase in multiple R does not reach statistical significance beyond the first step, and even with all the variables in the equation only 8.4 per cent of the variance is explained.

TABLE 4
SUMMARY OF THE FIRST FOUR STEPS OF THE REGRESSION OF ATTITUDE TO
SECONDARY SCHOOL WITH ALL THE PRIMARY SCHOOL VARIABLES

Primary School Variables	Multiple R	R^2	R^2 Change
1. Sex	0.17591*	0.03094	0.03094
2. Teacher's predictive rating	0.22566	0.05092	0.01998
3. Lie score	0.24719	0.06110	0.01018
4. Reading quotient	0.26348	0.06942	0.00832

$P < 0.05$

Table 5 summarises the first four steps of the regression analysis of secondary school attendance with all the primary variables. Attendance in the primary school accounts for 47.3 per cent of the variance with the rest of the variables explaining only a further 1.8 per cent of the total variance accounted for.

TABLE 5
SUMMARY OF THE FIRST FOUR STEPS OF THE REGRESSION OF ATTENDANCE IN
THE SECONDARY SCHOOL WITH ALL THE PRIMARY SCHOOL VARIABLES

Primary School Variables	Multiple R	R^2	R^2 Change
1. Attendance	0.68766*	0.47287	0.47287
2. Verbal reasoning quotient	0.69866*	0.48813	0.01526
3. Teacher's predictive rating	0.69957	0.48940	0.00126
4. Behaviour score	0.70027	0.49038	0.00098

* $P < 0.05$

DISCUSSION AND CONCLUSIONS

As expected, the higher correlations were obtained between those variables that doubled as primary school and criterion variables, i.e., teacher's rating, behaviour score and attendance. However, with the exception of attendance, the measure of agreement was low. Whilst this lack of agreement may partly reflect different teachers having different perceptions or employing somewhat different criteria in judging children, as well as a degree of unreliability in the measuring instruments themselves, it seems likely that there were genuine changes in many of the children's behaviour after transfer. There are three probable explanations for these changes. In some cases the change of behaviour may have been caused primarily by the traumatic or stimulating effect of the transfer as such. In other cases the change in behaviour was probably just a different reaction to a different situation (Coleman *et al.*, 1977). There is also the probability that life events, other than the transfer, during the follow-up period, would have affected some of the children's behaviour. Obviously, interacting combinations of these factors would apply in many cases.

The multiple correlation of 0.50 between the nine primary school variables and the behaviour score (Ruttler Scale B(2)) compares closely with the 0.52 which Pumfrey and Ward (1976) obtained using the Bristol Social Adjustment Guide as criterion. In that study there were also nine predictor variables which were similar to those in the present study. However, there was a four-year follow-up between primary and secondary school (10-14 years). Although there were also other methodological differences,

the similarity of the coefficients suggests that, in Pumfrey and Ward's study, a high proportion of the changes in the children's adjustment probably occurred in the first year of the follow-up period during which the children transferred schools. As in the present study, Pumfrey and Ward found that the initial measure of behaviour was the best predictor and that little was gained by including more than a few independent variables.

The present findings indicate that variables that are easily measured in the primary school have only a weak relationship with various criteria of adjustment in the secondary school. Only school attendance rates show real consistency across the transfer. Whereas other workers, e.g., Nisbet and Entwistle (1969), found that intellectual ability in the primary school was the main determinant of achievement in the secondary school, intellectual measures in this study have only low correlations with non-intellectual measures after transfer. The particularly weak relationship between the 'attitude to school' measure and the primary variables casts some doubt on the validity of the Fitt scale. However, as an index of adjustment to school, attitude appears to be quite different from the other criteria in that it reflects, to some extent, the child's own perception of his adjustment to school. Hawthorn (1976) and Merton (1975) found little agreement between teachers' assessments of children's adjustment to school and the children's self assessment. Therefore, the present results regarding attitude to school could be interpreted as consistent with these findings.

The results suggest that forecasts by primary school teachers, either on their own or used in conjunction with psychometric or educational tests, are likely to be rather inaccurate. This implies that secondary teachers should be very cautious in attempting to deduce from primary school records the probable level of children's adjustment in their first year after transfer.

ACKNOWLEDGMENTS.—The author would like to thank Professor Ronald Davie for the advice given to him in this study. He is also grateful to all the teachers whose co-operation made the investigation possible.

REFERENCES

- BALL, P. (1975). Some problems of children at transfer from junior to secondary school. In Educational Development Centre, *Continuity in Education*. Birmingham: City of Birmingham Education Department.
- BLYTH, W. A. L. (1965). *English Primary Education*. London: Routledge and Kegan Paul.
- BOWER, E. M. (1969). *Early Identification of Emotionally Handicapped Children in School*. (2nd edit.) Springfield, Ill.: Thomas.
- COLEMAN, J., WOLKIND, S., and ASHLEY, L. (1977). Symptoms of behaviour disturbance and adjustment to school. *J. child Psychol. Psychiat.*, 18, 201-209.
- COWEN, E. L., PEDERSON, A., BADIGIAN, H., IZZO, L. D., and TROST, M. A. (1973). Long term follow-up of early detected vulnerable children. *J. consult. clin. Psychol.*, 41, 438-446.
- COX, T. (1978). Children's adjustment to school over six years. *J. child Psychol. Psychiat.*, 19, 363-371.
- DAVIE, R., BUTLER, N., and GOLDSTEIN, H. (1972). *From Birth to Seven*. London: Longman.
- DOUGLAS, J. W. B. (1967). *The Home and the School*. St. Albans: Panther.
- EYSENCK, S. B. G. (1965). *Manual of the Junior Eysenck Personality Inventory*. London: University of London Press.
- FITT, A. B. (1956). An experimental study of children's attitude to school in Auckland, N.Z. *Br. J. educ. Psychol.*, 26, 25-30.
- GRONLUND, N. E., and HOLMLUND, W. J. (1958). The value of elementary school sociometric status scores for predicting pupils' adjustment in high school. *Educ. Adm. Superv.*, 44, 255-260.
- HALSALL, E. (1975). *The Comprehensive School*. London: Routledge and Kegan Paul.
- HAWTHORN, D. (1976). Transfer from primary to secondary education. Unpublished Dip. Ed. dissertation. University College, Cardiff.
- LIEM, G. R., YELLOTT, A. W., COWEN, E. L., TROST, M. A., and IZZO, L. D. (1969). Some correlates of early detected emotional dysfunction in the schools. *Am. J. Orthopsychiat.*, 39, 619-626.
- MERTON, J. (1975). Transfer from primary to secondary education. Problems of adjustment. In Educational Development Centre, *Continuity in Education*. Birmingham: City of Birmingham Education Department.

- NATIONAL FOUNDATION FOR EDUCATIONAL RESEARCH IN ENGLAND AND WALES (1971). Reading test SRB. Slough: NFER.
- NISBET, J. D., and ENTWISTLE, N. J. (1969). *The Transition to Secondary Education*. London: University of London Press.
- PIMM, J. B., and MCCLURE, G. (1967). A screening device for early detection of emotional disturbance in a public school setting. *Except. Child.*, 33, 647-648.
- PLOWDEN COMMITTEE (1967). *Children in Their Primary Schools. Vol. 1*. London: HMSO
- PUMFREY, P. D., and WARD, J. (1976). Adjustment from primary to secondary school. *Educ. Res.*, 19, 25-34.
- RUTTER, M. (1967). A children's behaviour questionnaire for completion by teachers: preliminary findings. *J. child Psychol. Psychiat.*, 8, 1-11.
- ULLMAN, C. A. (1952). Identification of maladjusted school children. *Public Health Monogr.*, No. 7, Public Health Service Publication No. 211. Washington, D.C.: Federal Security Agency.
- ULLMAN, C. A. (1957). Teachers, peers and tests as predictors of adjustment. *J. educ. Psychol.*, 48, 257-267.
- UNIVERSITY OF EDINBURGH (1974). *Manual for Moray House Verbal Reasoning Test 91*. London: University of London Press.
- WEDGE, P., and PROSSER, H. (1973). *Born to Fail?* London: Arrow.
- WESTMAN, J. C., RICE, D. L., and BERMAN, E. (1967). Nursery school behaviour and later school adjustment. *Am. J. Orthopsychiat.*, 37, 725-731.
- YELLOTT, A. W., LIEM, G. R., and COWEN, E. L. (1969). Relationships among measures of adjustment, sociometric status and achievement in third graders. *Psychol. Sch.*, 6, 315-321.
- YOUNGMAN, M. B. (1978). Six reactions to school transfer. *Br. J. educ. Psychol.*, 48, 280-289.

(Manuscript received 24th April, 1979)

INTERACTION AMONG DIMENSIONS OF ACADEMIC MOTIVATION AND CLASSROOM SOCIAL CLIMATE: A STUDY OF THE PERCEPTIONS OF JUNIOR HIGH AND HIGH SCHOOL PUPILS

By P. S. FRY*

(*The University of Calgary*)

AND

K. J. COE

(*Calgary General Hospital*)

SUMMARY. Previous studies have found a relationship between classroom social climate dimensions and specific academic outcome factors. The present study explored relationships among multidimensional characteristics of classroom social climates and pupil academic motivations in those settings. Classrooms perceived to be high in teacher support and involvement were associated with student motivations of self-improvement, academic success and enjoyment of learning. By contrast, classrooms perceived to be teacher-controlled or competition-oriented were related with anti-school feelings and a relative absence of self-improvement desires and enjoyment of learning. Task-oriented classrooms were associated with pupils expressing a greater desire for career preparation. Findings of these relationships are of significance for social systems intervention. Implications of these relationships are discussed in terms of reducing dissonance between classroom influences and negative affects of pupils, and for achieving greater congruence between positive motivations and classroom climate.

INTRODUCTION

SEVERAL recent investigations (e.g., Stern, 1970; Trickett and Moos, 1973, 1974; Randhawa and Michayluk, 1975; Moos, 1978; Moos and Moos, 1978) have emphasised the importance of assessing the impact of the social environment of the classroom on the social and emotional functioning of students within the classrooms. The underlying assumptions of this line of research are that (a) the environment of the classroom is defined by the shared perceptions of the members of the classroom along a number of environmental 'dimensions' (Trickett and Moos, 1974); (b) that environments exercise important psychological influence over their members (Nielsen and Kirk, 1974); and (c) these latter influences mediate the academic outcomes and motivations of the members within the environments (Moos, 1974; Epstein and McPartland, 1976; Maehr, 1976).

Trickett and Moos (1974) used a Classroom Environment Scale to link student satisfaction and mood to the social environment of the classroom; and Moos and Moos (1978) showed that the classroom social climate was significantly related to academic motivational factors of student grades and student absenteeism. As observed by these authors, classrooms characterised by a social climate of innovation are not only cohesive and satisfying but are intellectually challenging and encourage academic motivation of students. The findings of Manderscheid *et al.* (1977) support the notion that classroom psychosocial environments reflect the motivational dynamics that occur in that setting. Such effects have become a subject of continuing theoretical and empirical interest to researchers (e.g., Wicker, 1972; Jessor and Jessor, 1973; Randhawa and Fu, 1973) who have presented a theoretical analyses of the linkage of classroom environment and behavioural consequences and motivations.

* Requests for reprints should be sent to: P. S. Fry, Department of Educational Psychology, The University of Calgary, Calgary, Alberta, Canada T2N 1N4.

At present we know very little about why differences in academic motivations emerge in the particular form in which they do: for example, why specifically do certain pupils in certain classrooms become more 'anti-school' or more concerned about 'school grades' or 'self improvement'? As proposed by various authors (e.g., Anderson and Walberg, 1974; Hearn and Moos, 1978) proposing and testing concepts of subject-based, career-based, personality-based and motivation-based variations in classroom climates can lead to a more precise understanding of the empirically observed patterns of associations and relationships. Such assessments of relationships between social climates of the classroom and pupil reactions, interests and motivations may not only provide feedback to the teacher and taught about the characteristics of their milieu, but also motivate them to change the social milieu in ways that are self-enhancing to their own goals and aspirations (Moos, 1974). In the light of these discussions, it seemed reasonable to conjecture that various factors of academic motivation are related to the social climate of classrooms. Thus, for example, it is quite conceivable that highly competitive classrooms with high task orientations may be breeding grounds for motivations of academic success and grade-consciousness in pupils, while innovative classroom climates might lead to 'enjoyment of learning'. Similarly, there is reason to conjecture that social climates of classrooms perceived to be characterised by high teacher control may be associated with a passive academic orientation, and anti-school motivation; whereas, classroom environments perceived to be high in teacher support, involvement and affiliation might produce a social climate that is related to high pupil motivation for self-improvement and assertive participation in academic pursuits.

Overall, as suggested by Moos and Moos (1978), classroom environments are assumed to have certain demand characteristics which condition pupil motivations and manipulate their development along specific dimensions. Recognising the importance of understanding the academic motivations of students and recognising also the influence of classroom social climates in enhancing or inhibiting academic motivations, the present study was designed to examine interrelationships among specific characteristics of the classroom environment and the kinds of academic motivations that may prevail within different social climates.

The present study was further predicated on the assumption that academic motivations and goals of pupils are in large measure acquired through environmental conditioning and are not necessarily intrinsic and inherent in nature. Thus it seemed reasonable, as a first step, to study the interrelationships among the multivariable environmental dimensions of classrooms and academic motivational factors among pupils. As implied by Tuckman *et al.* (1969), such a multivariable assessment seems especially relevant at a time when student social systems seem to be significantly more change-oriented and students seem keen to direct, monitor and evaluate the environmental change process in ways that are self-serving and self-enhancing.

The present study used the academic motivation factors categorised by Doyle and Moen in a number of recent studies (e.g., Doyle and Moen, 1978; Moen, 1977; Moen and Doyle, 1977). Sociopsychological climate factors of the classroom were assessed by means of the Classroom Environment Scale (CES) which has also been used in a few recent investigations (e.g., Trickett and Moos, 1973; Moos and Trickett, 1974; Moos and Moos, 1978). The objective was to establish links between the social climate dimension of classrooms and the motivational and academic strivings of pupils within the classrooms.

METHODS

The study was carried out with students in 20 junior high and high schools. These schools represented a wide range of schools, e.g., middle class white schools, medium-sized rural white schools, small private schools, vocational high schools in

urban and rural areas. Sixty classrooms were sampled with three classrooms from each of the 20 schools. In most cases, classrooms were randomly sampled within schools: however in three schools a strict random sample was not possible because three classrooms randomly selected at first were involved in extra-curricular projects, and in three other classrooms the teachers were not anxious to participate. None of the classes was team taught. The average size of the class was 20 pupils. In seven classrooms with larger enrolment ranging between 35 and 45 pupils, 20 pupils were randomly selected to participate. Most classes included both male and female pupils, although there were a small number of all-male and all-female classes. In the high school sample, the pupils were mainly juniors and seniors and about one-fifth of the classes were being taken as electives. Various types of subject matter were represented, including English, social studies, biology, physical sciences, mathematics, foreign languages, business and technical classes. Thus the sample included a wide variety of classrooms designed to represent the range currently found in junior high and high schools.

Assessment of classroom climate

The Classroom Environment Scale (CES) (Trickett and Moos, 1974) which assesses classroom environments as perceived by students was used to assess classroom social climate. The underlying assumption in this scale is that the consensus of individuals characterising an environment constitutes a measure of the social climate of that environment. The CES, as described in Hearn and Moos (1978) and Moos (1978), consists of 90 true-false items which fall into nine different subscales, each of which measures the emphasis on one dimension of classroom climate. The involvement, affiliation, and teacher support dimensions relate to personal affective teacher-pupil or pupil-pupil relationships. A second group of subscales assesses the degree to which the class is task-oriented and focuses on competition. The last four subscales of order and organisation, rule clarity, teacher control and innovation tap information about maintenance and authority functions relevant to the structure and organisation of the classroom, as well as about the processes and potential changes in classroom functioning. Further details about the development and correlations of the CES are provided in Trickett and Moos (1973) and Moos and Trickett (1974). Briefly, the subscales measure distinct aspects and orientations of classroom environments. Each of the subscales significantly discriminates among classrooms, and internal consistencies and profile stability are quite high (Hearn and Moos, 1978).

The CES was administered to ten pupils randomly selected within each of the 60 classrooms. Mean scores (based on the CES scores of ten pupils) for each of the subscales were then obtained for each classroom. Classroom standard scores for each subscale were obtained by comparing the mean subscale score for each classroom to the overall mean of 60 classrooms.

Assessment of academic motivations of pupils

The Academic Motivations Inventory (AMI) was used to assess academic motivations of students in the classroom (Moen and Doyle, 1977; Doyle and Moen, 1978). This is an 80-item self-report instrument. Each item is cast as a simple declarative sentence on a five-point Likert-type scale: 'Not at all true of me' through 'Extremely true of me'. Academic motivations are defined by the authors of the AMI scale as those material, psychological, or social forces that draw or impel students to put effort into curricular activities. The authors of the AMI have identified nine scales or areas of motivation for which Hoyt internal consistency coefficients range from 0.62 to 0.83. Immediate retest estimates of reliabilities for the scales range from 0.86 to 0.93. The nine factors of Academic Motivation identified by the authors of the AMI (Doyle and Moen, 1978) are Desire for Self-Improvement (15 items), Anti-School (17 items), Desire for Esteem (14 items), Enjoyment of Learning

(11 items), Enjoyment of Assertive Interactions (seven items), Resentment of Poor Teaching (three items), Academic Success (four items), Desire for Career Preparation (four items), and Enjoyment of Passive Interactions (five items). Each of these subscales significantly discriminates among groups of students.

The AMI was administered to ten pupils randomly selected within each of the 60 classrooms. Mean scores (based on the AMI scores of ten pupils) for each of the nine subscales were then obtained for each classroom. Classroom standard scores for each subscale were obtained by comparing the mean subscale score for each classroom to the overall mean of 60 classrooms.

It should be noted that while the AMI and CES were administered to all classrooms within the 20 schools selected for the study, the CES scores and AMI scores for a particular classroom were derived from the responses of different individuals. In 30 classrooms, ten pupils were randomly selected from the sample within each class to take the CES with the remaining ten pupils required to take the AMI. In the 30 other classrooms, the procedure was reversed and ten pupils randomly selected from the sample of the class were administered the AMI and the remaining ten pupils took the CES. Such a procedure was designed to ensure that the resulting correlations between pupils' scores for academic motivation and classroom environment were not simply a consequence of the individual pupil's rating set but that they reflected a more general and overall relationship between pupils' perceptions of the classroom environment dimensions and pupils' self-reports of academic motivation factors. Both the AMI and CES were administered to the pupils by project assistants in the absence of the teacher because previous studies had reported anxiety in the presence of instructors.

Pupils provided information about themselves and the climate of the classrooms in both the AMI and the CES and pupil self-perceptions and self-reports were the major focus of the study. As Stern (1970) points out, there may be some disparity between the perceived situations of social climate and academic motivations and the veridical ones; however, for the pupils themselves, the perception is the reality.

RESULTS

In order to ensure that individual pupils in the classroom agreed on the characteristics of the classroom as portrayed by the subscale dimensions of the CES, an intercorrelation matrix of the individual pupil scores and classroom scores was examined (see Table 1). Similarly, in regard to the subscale dimensions of pupils' academic motivations, an intercorrelation matrix of individual pupil scores and classroom scores was examined (see Table 2). Intercorrelation matrices for both the CES and AMI showed a consistently high degree of agreement among individual pupil scores and classroom scores, thus suggesting that scores for each classroom could now

TABLE 1
INTERCORRELATIONS OF INDIVIDUAL PUPIL SCORES AND CLASSROOM SCORES
ON SUBSCALES OF THE CES

CES Classroom scores (N = 60) with CES Individual pupils (N = 600)									
CES Subscales	Inv.	Aff.	T.S.	T.O.	Comp.	Org.	R.C.	T.C.	Inn.
	0.83**	0.82**	0.84**	0.57*	0.81**	0.71**	0.53*	0.81**	0.74**

*P < 0.05

** P < 0.01

Inv. = Involvement; Aff. = Affiliation; T.S. = Teacher Support; T.O. = Task Orientation; Comp. = Competition; Org. = Organisation and Order; R.C. = Rule Clarity; T.C. = Teacher Control; Inn. = Innovation.

TABLE 2

INTERCORRELATIONS OF INDIVIDUAL PUPIL SCORES AND CLASSROOM SCORES
ON SUBSCALES OF THE AMI.

AMI Classroom scores (N = 60) with AMI Individual pupils (N = 600)									
AMI Subscales	S.I. Des.	Anti-Sch.	Est. Des.	En. Lng.	Ass. In't	Resmt. Poor Tchg.	Acad. Succ.	Car. Prep. Des.	Des. Pass. In't.
	0.87**	0.71**	0.85**	0.76**	0.72**	0.63*	0.86**	0.59*	0.58*

* PM0-05

** PM0-01

S.I. Des. = Self Improvement Desire; Anti Sch. = Anti-School; Est. Des. = Desire for Esteem; En. Lng. = Enjoyment of Learning; Ass. In't = Assertive Interactions; Resmt. Poor Tchg. = Resentment of Poor Teaching; Acad. Succ. = Academic Success; Car. Prep. Des. = Desire for Career Preparation; Des. Pass. In't. = Desire for Passive Instruction.

be combined to yield an average score for the classrooms on the subscales of the CES and AMI. Subsequently for the 60 classrooms, the resulting nine subscale scores for the nine dimensions of the CES were correlated with the resulting nine subscale scores for the nine dimensions of the AMI.

There were substantial relationships between pupil perceptions of the classroom environment dimension and pupil self-reports of academic motivation factors (see Table 3). Classrooms perceived to be characterised by high involvement and teacher

TABLE 3

CORRELATIONS OF CES SUBSCALE SCORES AND AMI SUBSCALE SCORES

CES Sub-scales	Desire for Improvement	Anti-school	Desire for Esteem	Enjoyment of Learning	Assertive Interactions	Resentment of Poor Teaching	Academic Success	Desire for Career Preparation	Desire for Passive Interaction
Inv.	0.62**	-0.48*	0.26	0.65**	0.15	0.10	0.67**	0.24	-0.21
Aff.	0.34	-0.36	0.36	0.21	0.10	-0.19	0.27	0.15	-0.16
T.S.	0.65**	-0.63**	0.35	0.45*	0.18	-0.15	0.45*	0.07	0.26
T.O.	0.18	-0.16	0.16	0.03	0.27	0.12	0.31	0.67**	-0.03
Comp.	-0.30	0.47*	0.26	-0.45*	0.35	0.18	0.35	0.03	-0.40*
Org.	0.16	-0.19	0.05	0.03	0.17	0.08	0.10	0.44†	-0.18
R.C.	0.12	0.03	0.10	0.05	0.26	0.15	0.03	0.16	0.05
T.C.	-0.52*	0.69**	-0.16	-0.45*	0.20	0.40	0.22	0.10	0.37
Inn.	0.18	-0.22	0.21	0.67**	0.35	0.45*	0.28	0.16	0.08

* P < 0.05

** P < 0.01

Inv. = Involvement; Aff. = Affiliation; T.S. = Teacher Support; T.O. = Task Orientation; Comp. = Competition; Org. = Organisation and Order; R.C. = Rule Clarity; T.C. = Teacher Control; Inn. = Innovation.

support held pupils with a higher academic motivation for self-improvement, enjoyment of learning and the desire for academic success. Involvement and teacher support were negatively correlated with Anti-School motivation. High Teacher Control was positively correlated with Anti-School feeling, and negatively correlated with 'enjoyment of learning'. Innovative practice in the classroom was positively correlated with Enjoyment of Learning. Task orientation and organisation and order dimensions of the classroom climate were positively correlated with a desire for career preparation.

Thus, the number of factors in the classroom climate which combined to a significant degree with factors of academic motivation were: Involvement and Teacher

TABLE 4

SPECIFIC ITEMS OF THE AMI SUBSCALES THAT ARE SIGNIFICANTLY RELATED TO CES CLASSROOM ORIENTATIONS: TEACHER CONTROL AND COMPETITION

Teacher Control: assesses how strict the teacher is in enforcing the rules, and the severity of the punishment for rule infractions.

Competition: assesses the emphasis placed on students competing with each other for grades and recognition.

		Items
<i>Teacher Control</i>	Correlation with Anti-School Feelings 0.69	—School gives me a feeling of insignificance (true) —I dislike most school work (true)
	Correlation with Enjoyment of Learning -0.45	—I really enjoy learning new things (false) —I enjoy reading books and materials teachers assign me (false).
	Correlation with Self-Improvement -0.52	—I hope school will make me feel able to handle whatever comes up (false).
<i>Competition</i>	Correlation with Anti-School Feelings 0.47	—Grades or degree requirements put restrictions on me that I hate (true) —School gives me a feeling of insignificance that I hate (true)
	Correlation with Enjoyment of Learning -0.45	—It is very important that my classes make me use my mind (false) —I just really enjoy learning new things (false).

All significant at greater than the 5% level.

Support, Competition, and Teacher Control, and to a somewhat lesser degree with factors of Innovation and Task Orientation.

On the other hand, factors of academic motivation which combined significantly with a number of factors of classroom climate were Desire for Self-Improvement, Anti-School feelings (or Negative Motivation towards School) Desire for Academic Success (or conscientiousness about grades, and school work), and Enjoyment of Learning (or Positive Affect and Positive Motivation towards School.)

In summary, analyses of classrooms in terms of specific characteristics show some significant relationships: classrooms perceived to be 'affiliation oriented' (i.e., characterised by emphasis on student-teacher support and interaction, cf. Moos, 1978) had students who reported a high desire for self-improvement, who enjoyed learning and who were motivated towards academic success. By contrast, classrooms perceived to be 'competition oriented' (Moos, 1978) had students reporting high anti-school feelings and lowered enjoyment of learning. Classrooms generally perceived to be 'control oriented' (Moos, 1978) (i.e., in which teacher control is high and the emphasis is also on organisation and order, cf. Moos, 1978) were associated with pupils reporting a significant degree of anti-school feeling and a lowered interest in self-improvement. Lastly, classrooms perceived by pupils to be 'task oriented' were associated significantly with students who were career minded and motivated by a desire for career preparation and specific task objectives.

Since teacher control, involvement, teacher support, and competition were

TABLE 5

SPECIFIC ITEMS OF THE AMI SUBSCALES THAT ARE SIGNIFICANTLY RELATED TO CES CLASSROOM ORIENTATIONS: TEACHER SUPPORT AND INVOLVEMENT.

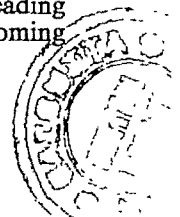
Teacher Support: assesses the amount of help, concern, and friendship the teacher directs towards students.

Involvement: assesses the extent to which students and teacher have attentive interest in class activities.

		Items
<i>Teacher Support</i>	Correlation with Academic Success 0.45	—I am enthusiastic about trying to get high scores or grades (true) —I try to do my very best on all my school work (true)
	Correlation with Enjoyment of Learning 0.45	—I enjoy reading most books and articles teachers assign to me (true)
	Correlation with Self-Improvement 0.65	—I hope school will help me become a better person (true)
	Correlation with Anti-School -0.63	—School gives me a feeling of insignificance that I hate (false)
<i>Involvement</i>	Correlation with Self-Improvement 0.62	—I hope school will teach me better ways of handling conflicts with people (true) —I hope school will help me become a better person (true)
	Correlation with Enjoyment of Learning 0.65	—I just really enjoy learning new things (true) —I really enjoy reading most books and articles my teachers assign (true)
	Correlation with Academic Success 0.67	—I try to do my very best on all my school work even on things uninteresting (true) —If I get started on something, it is very important for me to complete it (true)

All significant at greater than the 5% level.

dimensions of the classroom climate that seemed to be most closely and significantly associated with dimensions of pupil motivations, a further analysis was conducted to clarify these specific aspects of the classroom milieu and to relate these to specific items of the AMI motivation subscales. Each of the AMI items is answered either 'true of me' or 'not true of me' (false) along a five-point continuum. The proportion of pupils answering 'true' or 'false' was calculated for each item of the AMI in classrooms predominantly perceived to be 'Teacher controlled' (N = 15 classrooms); 'Competition Based' (N = 16 classrooms); and 'With Teacher Support' (N = 10 classrooms) and having 'Involvement' (N = 19 classrooms). Specific items of the AMI subscales that are significantly ($P < 0.05$) related to classroom orientations are shown in separate tables (see Tables 4 and 5). Pupils in classrooms with Teacher Control and Competition felt insignificant, disliked school work and felt restricted by grades. Pupils in classrooms with Teacher Support and Involvement enjoyed reading books, felt enthusiastic about grades, felt significant and concerned about becoming better individuals.



DISCUSSION

The most general rationale behind the present study was to provide data relevant to the proposition that different classroom environments and social climate of classrooms are systematically related to the academic motivational states of the pupils. The findings of the present study confirmed that there are distinct relationships between certain social climates of the classroom as perceived by the pupils, and certain dimensions of academic motivation. For example, there appears to be a distinct association between classrooms characterised by teacher support and involvement with pupil motivations of enjoyment of learning, a desire for academic success and a desire for self-improvement.

In addition to this general picture of the 'motivating' classroom, another area of interest involved 'anti-school' feelings or negative motivation. Because such a motivation is seen as a pervasive problem area for most classrooms and class teachers, the kind of classroom environment or social climate that relates to this problematic issue has important educational implications. The present data suggest that anti-school feelings predominate in classrooms perceived to have a climate of competition and restrictive teacher control. Thus, the emergent picture of the classroom where pupils report a competitive climate and high teacher control combines an affective concern of students with very limited enjoyment of learning. These findings about the motivational correlates of classroom environments are generally congruent with the Trickett and Moos (1974) data on pupil satisfactions and the 'satisfying' classroom in which there is personal pupil-teacher relationship and high involvement. Task orientation and teacher control in classrooms are often rationalised on the basis that pupils will strive for greater academic success. That this relationship was not found may have implications for educational ideology and practice. No causal implications can be drawn from the data concerning pupil motivations in classroom climates of competition and restrictive teacher control and those classrooms characterised to be interactive and high in teacher support. In fact, differing pupil motivations and classroom climates and specific dimensions of competition, teacher control and teacher support are probably interrelated in a complex manner. It would be hard to establish whether classroom climates of competition or teacher control or teacher support shape academic motivations of pupils or whether pupils come into classrooms with academic motivations which are a product of previous conditioning, and which in turn positively enhance or distort their perceptions of the classroom climate. Indeed, teachers are not enduringly or persistently 'controlling' or 'supportive', but may be affected by pupil reactions to become so.

Teachers who continually interact with pupils who demonstrate a desire for self-improvement, a love of learning and an ambition for academic success, can well afford to be supportive and involved with their pupils. They have relatively little need to justify their authority (Kaye *et al.*, 1977). By contrast, teachers who must contend with pupils who display anti-school feelings or who show little enjoyment of learning, are likely to become restrictive and controlling in class and apt to criticise pupils for their poor motivations. Similarly, pupils who enjoy learning and teachers who are innovative will likely experience a mutual attraction.

In sum, classroom social climates not only reflect the motivational dynamics of the participants in those settings, but are, in turn, shaped and manipulated by the motivational reactions that transpire there (Manderscheid *et al.*, 1977). By assessing both classroom environments and academic motivations of pupils in a multidimensional fashion, it is possible to determine the degree of congruence or dissonance that exists between a particular classroom climate and the motivational needs of the pupils. Social systems intervention can then focus on reducing the degree of dissonance or incompatibility that exists between a given group of pupils with specific motivations and the classroom climate to which they must adjust. For example, the present

findings suggest that decreasing competition and teacher control for pupils who have a desire for self-improvement, may create a social climate in which the pupil has opportunity for growth and development. Similarly, pupils who have 'anti-school' feelings or 'motivation away from school' are likely to feel more at ease in classroom environments which are relatively free of competition. Thus, social systems intervention that focuses on increasing social support and teacher involvement in settings in which anti-school feelings run high, is likely to produce more motivation-environment congruence. By contrast, an environment high in competition combined with students who have strong anti-school feelings will most likely increase motivation-environment dissonance.

Many educators will want to know how they can affect pupil motivation, by fostering change in the pupil, or change in the classroom environment. To determine how to change pupils' motivational characteristics or the motivational characteristics of the environment requires information about multivariable dimensions of classrooms and pupils motivations. Data from the CES and AMI can be used to assess the degree of dissonance or consonance that exists between classroom climates and pupil motivations. The information can then be used to formulate planned change in the environment so that it becomes compatible with the kinds of strivings or needs that pupils have. Compatibility of classroom climate and pupil needs and motivations will influence how hard pupils will work in school, how well they will wish to perform and how satisfied they will be with their educational experience (Moen and Doyle, 1978).

REFERENCES

- ANDERSON, G., and WALBERG, H. J. (1974). Learning environments. In WALBERG, H. J. (Ed.), *Evaluating Educational Performance: a Sourcebook of Methods, Instruments and Examples*. Berkeley, Calif.: McCutchan.
- DOYLE, K. O., and MOEN, R. E. (1978). Toward the definition of a domain of academic motivation. *J. educ. Psychol.*, 70, 231-236.
- EPSTEIN, J., and MCPARTLAND, J. (1976). The concept and measurement of the quality of school life. *Am. educ. Res. J.*, 13, 15-30.
- HEARN, J., and MOOS, R. (1978). Subject matter and social climate in high school classrooms: a test of Holland's theory. *Am. educ. Res. J.*, 15, 111-124.
- JESSOR, R., and JESSOR, S. L. (1973). The perceived environment in behavioral science. *Am. Behav. Scientist*, 16, 801-828.
- KAYE, S., TRICKETT, E., and QUINLAN, D. (1977). Alternative methods for environmental assessment: an example. *Am. J. Community Psychol.*, 4, 367-377.
- MAEHR, M. (1976). Continuing motivation: an analysis of a seldom considered educational outcome. *Rev. educ. Res.*, 46, 443-462.
- MANDERSCHIED, R. W., KOENIG, G. R., and SILBERGELD, S. (1977). American dimensions of classroom psychosocial environment. *Am. J. Community Psychol.*, 5, 299-306.
- MOEN, R. E. (1977). Toward the measurement of academic motivation. Unpublished doctoral dissertation, University of Minnesota, Minneapolis. Minnesota.
- MOEN, R., and DOYLE, K. O. (1977). Construction and development of the academic motivations inventory. *Educ. psychol. Measur.*, 13, 509-512.
- MOEN, R., and DOYLE, K. O. (1978). Measures of academic motivation: a conceptual review. *Res. in Higher Educ.*, 8, 1-23.
- MOOS, R. (1974). *Evaluating Treatment Environments: a Social Ecological Approach*. New York: Wiley Interscience.
- MOOS, R. (1978). Toward a typology of classroom social environments. *Am. educ. Res. J.*, 15, 53-60.
- MOOS, R., and MOOS, B. S. (1978). Classroom social climate and student absences and grades. *J. educ. Psychol.*, 70, 263-269.
- MOOS, R., and TRICKETT, E. (1974). *Classroom Environment Scale Manual*. Palo Alto, Calif.: Consulting Psychologists Press.
- NEILSON, H. D. and KIRK, D. (1974). Classroom climates. In WALBERG, H. (Ed.), *Evaluating Educational Performance*. Berkeley, Calif.: McCutchan.
- RANDHAWA, B., and FU, L. (1973). Assessment and effects of some classroom environment variables. *Rev. educ. Res.*, 43, 303-321.
- RANDHAWA, B., and MICHAYLUK, J. (1975). Learning environment in rural and urban classrooms. *Am. educ. Res. J.*, 12, 265-285.

- STERN, G. (1970), *People in Context: Measuring Person—Environment Congruence in Education and Industry*, New York: Wiley.
- TRICKETT, E., and MOOS, R. (1973). The social environment of junior high and high school classrooms. *J. educ. Psychol.*, 65, 93-102.
- TRICKETT, E., and MOOS, R. (1974). Personal correlates of contrasting environments: student satisfaction in high school classrooms. *Am. J. Community Psychol.*, 2, 1-12.
- TUCKMAN, B., MCCALL, K., and HYMAN, R. (1969). The modification of teacher behavior. Effects of dissonance and coded feedback. *Am. educ. Res. J.*, 6, 607-619.
- WICKER, A. W. (1972). Processes which mediate behavior-environment congruence. *Behav. Sci.* 17, 265-277.

(Manuscript received 27th March, 1979)

SOME DETERMINANTS OF EARLY SECONDARY SCHOOL PERFORMANCE

By M. B. YOUNGMAN

(School of Education, University of Nottingham)

SUMMARY. This paper reports further analyses of a study of primary/secondary transfer for which certain findings have already been published (Youngman and Lunzer, 1977; Youngman, 1978; Youngman, 1979). Here the concern is to examine the various pupil characteristics that might determine subsequent academic performance in secondary school. It employs covariance regression models rather than the more customary stepwise or complete models and consequently slightly different hypotheses are investigated. Specifically the analyses seek to determine what measures or sets of measures have independent effects on achievement or progress in reading or mathematics. Two samples ($N=290$ and 359) are analysed separately to test the possible generality of the results obtained.

The results presented comprise the two sample correlation matrices, a covariance regression summary table for reading achievement, and summary charts for all four performance criteria analyses. Most of the discussion and interpretation is based on these summary charts. From them it is seen that in spite of differences in intellectual level (Youngman, 1978) variations between the results for the two samples are small. The predominant pattern indicates that neither biographic nor dispositional (that is, attitude or personality) characteristics show any marked independent effect on performance. On the other hand intellectual characteristics show strong effects, both individually and in association with other intellectual measures.

INTRODUCTION

TRANSFER from primary to secondary education should not be considered a special feature of schooling since all educational experience involves the pupil in some kind of adjustment to new situations or demands. Nevertheless it does represent an important stage in education, if only because of the strength and variety of the influences pupils encounter during that period. In order to contain discussion it is convenient to categorise the various influences into situational (school characteristics mainly), biographic (pupil characteristics such as age, sex and social class), intellectual (ability and achievement) and dispositional (covering attitude and personality characteristics). Situational factors have been examined in a range of studies (e.g., Nisbet and Entwistle, 1966; Neal, 1975; Sumner and Bradley, 1977) but few recurrent themes or findings emerge. The other factors are predominantly pupil-orientated and seem to have generated issues of a more general nature. Particularly prevalent are studies of pupil adjustment, defined either as personal adjustment (e.g., Thompson, 1975; Pumfrey and Ward, 1976) or as academic adjustment (Dale and Griffith, 1965; Nisbet and Entwistle, 1969). The present study is directed towards the latter topic, the identification of factors influencing academic adjustment as measured by performance in basic subjects in the early secondary years.

In an earlier report (Youngman, 1978) it was shown that dispositional and intellectual characteristics were associated, although no attempt was then made to isolate specific influences on performance. Of studies that have, numerous individual characteristics have been shown to relate to academic performance. Nisbet and Entwistle (1969) established significant relationships with performance for various intellectual variables, and for certain dispositional variables including academic motivation, attitude to work and neuroticism. Behrens and Vernon (1978) showed that attitude to school, self-concept and aggression (all classified here as dispositional) were related to achievement. Biographic effects tend to be weaker but both the above studies identified significant relationships in this category. Nisbet and Entwistle

reported a social class effect, Behrens and Vernon one due to sex. With such a wide range of associations it becomes important to isolate the substantial determinants of performance. In particular, which characteristics influence performance independently, and not through their relationships with other, stronger determinants?

METHOD

Samples and measures

Earlier publications (Youngman and Lunzer, 1977; Youngman, 1978; Youngman, 1979) give details of the data for this study and anyone requiring more than a summary is referred to those sources. Briefly, two samples (a rural one of 290 pupils and a city one comprising 359 pupils) were assessed on a selection of biographic, dispositional and intellectual measures before and after transfer to secondary comprehensive schools. The secondary dispositional measures were applied after two terms in the school whilst the two performance assessments were made at the end of the second year. In addition to these two achievement scores, two corresponding progress measures were computed by constructing regression residuals, the difference between the achievement level *predicted* from the junior school score and the *actual* secondary school achievement level. Hence the complete analysis variable set is as follows:

<i>Variable</i>	<i>Code</i>	<i>Description</i>
BIOGRAPHIC		
3	AGE	Age
4	SEX	Sex (1 = male/2 = female)
DISPOSITIONAL		
6	JASS	Junior attitude to secondary school
7	JAPS	Junior attitude to primary school
8	JAPH	Junior apprehension over transfer
9	JSCS	Junior self concept: social
10	JSCP	Junior self concept: personal
11	JSCA	Junior self concept: academic
12	JMOT	Junior academic motivation
13	SASS	Secondary attitude to secondary school
14	SAPS	Secondary attitude to primary school
15	SANX	Secondary anxiety
16	SSCS	Secondary self concept: social
17	SSCP	Secondary self concept: personal
18	SSCA	Secondary self concept: academic
19	SMOT	Secondary academic motivation
INTELLECTUAL		
20	JNVR	Junior non-verbal reasoning
21	JRED	Junior reading comprehension
22	JMAT	Junior mathematical understanding
23	SRED	Secondary reading comprehension
24	SMAT	Secondary mathematical understanding
25	PRED	Progress in reading
26	PMAT	Progress in mathematics

Analyses

Following construction of the product-moment correlation matrix for each sample the two progress measures were computed using linear regression models predicting secondary achievement (SRED and SMAT) from junior school achievement (JRED and JMAT). For each pupil a progress score was defined as the difference between the *actual* secondary score and the *predicted* secondary score. This value is usually referred to as the residual score. The main advantage of this measure over a

simple difference score is that the residual is uncorrelated with the initial score. The simple difference is negatively correlated and therefore is biased in favour of low initial scores.

The main analyses were performed using covariance regression models. A multiple linear regression model comprises a criterion measure and a set of predictor measures. The computational procedure (here Veldman's, 1967, as presented in Youngman, 1976) seeks an optimal linear combination of the predictor variables such that this combined score correlates maximally with the criterion measure. This correlation, the multiple correlation R , has the value 1.0 for perfect prediction. As with the bivariate correlation coefficient the square of R is the proportion of the criterion variance predicted or explained by the predictor set, the model. Having established the predictive power of the model, it is then normal to examine the relative contributions of the individual components. Considerable dissension exists over how this can best be done, if at all, and Cooley and Lohnes (1971) give a good summary of the arguments. They favour the structure correlation, that is the correlation between a predictor and the predicted score. However as its formula shows,

$$r_{si} = \frac{1}{R} r_{ci}$$

where r_{ci} is the correlation between variable i and the criterion,
 R is the multiple correlation for the model,

the structure correlation does not take into account inter-correlation between predictor variables. For the purpose of the present study this interdependence must be accommodated within the analytical procedure.

One solution is to use analysis of covariance models. First a *full* multiple linear regression model is computed including all the variables to be evaluated in relation to a particular criterion. Then a *reduced* model is constructed *omitting* the variable (or set of variables) whose contribution is to be examined. If the multiple correlation drops substantially then the omitted variable contributes to the criterion variance *even when all the other predictors are present*. It can be said to have an independent effect. Clearly it is not sufficient simply to inspect the change in R . The statistical significance of the reduction is tested using an F -test:

$$F = \frac{(R_F^2 - R_R^2)/(K_F - K_R)}{(1 - R_F^2)/(N - K_F)}$$

where R_F and R_R are the full and reduced multiple correlations,
 K_F and K_R are the numbers of independent predictors per model.

The degrees of freedom for F are $(N - K_F - 1)$ for the denominator, $(K_F - K_R)$ for the numerator. One final comment concerns the difference between statistical significance and educational significance. Because the F -test operates on R^2 , for high values of R a very small drop can be statistically significant, particularly if the effect of omitting a single predictor is being tested. It is therefore important to examine the *amount* of the change in R^2 as well as its statistical significance.

The covariance model is very effective for examining the independent contributions of variables and sets of variables. For example it enables what are virtually test-retest combinations (such as junior academic motivation and secondary academic motivation) to be omitted *in toto*, avoiding the danger of hiding an effect by retaining an equivalent measure. Altogether 20 different effects were tested, covering various combinations of biographic, dispositional and intellectual variables. Four criteria were used; they were secondary *achievement* and *progress* in *reading* and *mathematics*.

RESULTS

Table 1 gives the product-moment correlations for the 23 variables used in the later analyses. Both samples are included, the rural sample values occupy the bottom left half of the matrix. The high values in the bottom right-hand corner of the table reflect the expected strong associations between intellectual measures. Even so, significant associations with performance are not restricted to other intellectual measures. Most of the dispositional measures (variables 6 to 19) show significant correlations with at least one of the intellectual measures, some of the values exceeding 0.3. Correlations between the biographic variables and performance are, however, uniformly low.

Table 2 follows Cooley and Lohnes' recommendation by reporting the structure correlations along with multiple correlations. It brings together 20 different reduced models using reading achievement as the criterion. The empty cells indicate which variables have been omitted and the covariance column gives the significance test result comparing the reduced model with the full model. Correlations between the individual predictor variables and the criterion are supplied below the table. Comparing this row with any model row confirms that the structure correlations are simply the criterion correlations multiplied by the factor $1/R$. It follows that to report structure correlations for every model, as in Table 2, is largely superfluous; if required they can quickly be obtained by dividing the relevant row of the correlation matrix by R . Consequently for the three remaining performance criteria, mathematics achievement, reading progress and mathematics progress, this complete table is not presented although it is available on request. Instead summary charts are compiled showing the relevant results in diagrammatic form. Interpretation is also related to these charts.

Figure 1 records the model results for both reading criteria with both samples. The squared multiple correlation is represented, rather than R , since it relates more closely to the significance tests. It also becomes a simple matter to represent R squared, the proportion of variance explained, as a bar of appropriate length. This appears as the speckled section. Reduced models are shown similarly, with the black portion representing the difference between the full and reduced models. More importantly it now becomes possible to qualify the F -test result by examining the *size* of the effect as well as its statistical significance. Other features of Figure 1 are relatively straightforward. The left and right hand blocks record the rural and city sample results respectively. Within each horizontal bar the upper half represents the achievement measure whilst the bottom represents progress.

Once the covariance results have been converted into diagrammatic form their interpretation is clearer. The significance markers indicate that statistically significant effects are associated predominantly with intellectual measures, the only exceptions being an age effect on reading progress for the rural sample (models 1 and 3) and a personality effect (secondary self concept and motivation, models 8, 9, 11, 12 and 13) on both achievement and progress for the city sample. However, when the size of the effect is also considered, then even in combination (for example, model 13) non-intellectual measures have small influences, variance differences rarely exceeding 5 per cent.

Before examining these results in greater detail, it is worth looking at the equivalent results for mathematics. In Figure 2 the distinction between intellectual and non-intellectual effects is even more marked. Only four marginal biographic effects break the pattern associating significance with intellectual predictors. Taken together Figures 1 and 2 show consistency in the way significant *and* substantial effects are located within the intellectual domain. The absence of significance for certain models (specifically with the progress measures of Figure 1, model 15, and Figure 2, model 16)

TABLE 2

COVARIANCE REGRESSION RESULTS FOR READING ACHIEVEMENT: BOTH SAMPLES

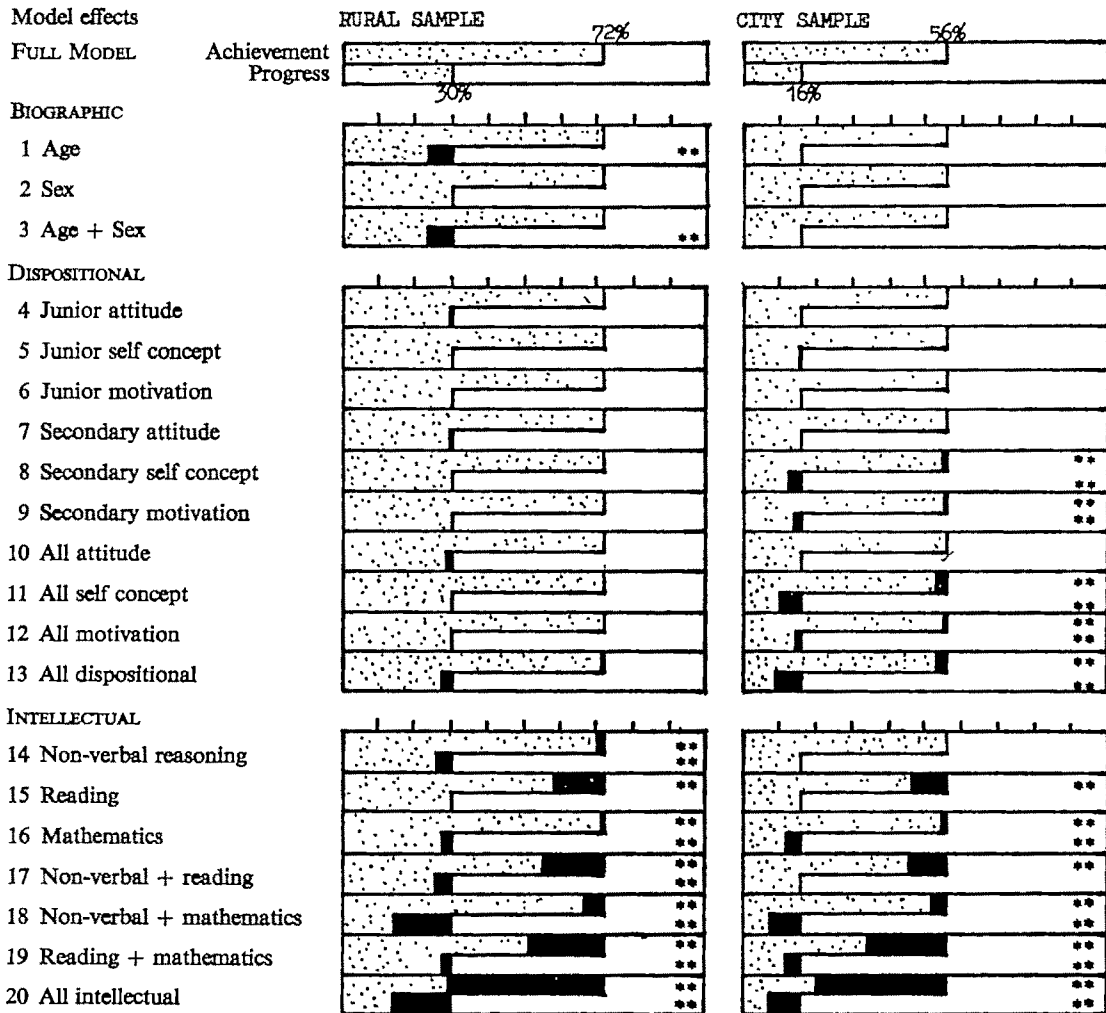
Model	Predictor variables with structure correlations in cells																						cov. R test
	3	4	6	7	8	13	14	15	9	10	11	16	17	18	12	19	20	21	22				
	A G E	S E X	A S S	J A P S	J A P H	J A S S	S A P S	S A P S	S A N X	J S C S	J S C P	J S C A	S S C S	S S C P	S S C A	J M O T	S M O T	J N V R	J R E D	J M A T			
Rural	-03	02	08	22	-31	12	00	-36	11	36	39	09	30	36	38	35	76	93	83			0.851	
City	09	18	06	32	-16	14	00	-16	05	37	25	-07	07	17	39	35	61	92	81			0.751	
1		02	08	22	-31	12	00	-36	11	36	39	09	30	36	39	35	76	93	83			0.851	
		18	06	32	-16	14	00	-16	05	37	25	-07	07	17	39	35	61	92	81			0.751	
2	-03		08	22	-31	12	00	-36	11	36	40	09	30	36	39	35	76	93	84			0.850	
	09		06	32	-16	14	00	-16	05	37	25	-07	07	17	39	35	61	92	81			0.751	
3			08	22	-31	12	00	-36	11	36	40	09	30	36	39	35	76	93	84			0.850	
			06	32	-16	14	00	-16	05	37	25	-07	07	17	39	35	61	92	81			0.749	
4	-03	02				12	00	-36	11	37	40	09	30	36	39	35	76	93	84			0.849	
	09	18				14	00	-16	05	37	25	-07	07	17	39	35	61	92	81			0.750	
5	-03	02	08	22	-31	12	00	-36				09	30	36	39	35	76	93	84			0.850	
	09	18	06	32	-16	14	00	-16				-07	07	17	39	36	62	93	81			0.747	
6	-03	02	08	22	-31	12	00	-36	11	36	39	09	30	36		35	76	93	83			0.851	
	09	18	06	31	-16	14	00	-16	05	37	25	-07	07	17		35	62	92	81			0.751	
7	-03	02	09	22	-31				11	37	40	09	30	36	39	35	76	93	84			0.848	
	09	18	06	32	-16				05	37	25	-07	07	17	39	35	61	92	81			0.750	
8	-03	02	09	22	-31	12	00	-36	11	36	39				39	35	76	93	83			0.851	
	09	19	06	32	-17	14	00	-16	05	38	25				39	36	63	94	82			0.736**	
9	-03	02	08	22	-31	12	00	-36	11	36	39	09	30	36	39		76	93	83			0.851	
	09	19	06	32	-16	14	00	-16	05	37	25	-07	07	17	39		62	93	81			0.744**	
10	-03	02							11	37	40	09	30	36	39	35	76	93	84			0.846	
	09	18							05	37	25	-07	07	17	39	35	61	93	81			0.749	
11	-03	02	08	22	-31	12	00	-36							39	35	76	93	84			0.850	
	10	17	06	32	-17	14	00	-16							40	36	63	95	83			0.731**	
12	-03	02	08	22	-31	12	00	-36	11	36	39	09	30	36			76	93	83			0.851	
	09	18	06	32	-16	14	00	-16	05	37	25	-07	07	17			62	93	81			0.743**	
13	-03	02															76	94	84			0.842	
	10	19															63	96	84			0.725**	
14	-03	02	08	22	-31	13	00	-37	11	37	40	09	30	36	39	36		94	85			0.838**	
	09	18	06	32	-16	14	00	-16	05	37	25	-07	07	17	39	35		92	81			0.750	
15	-03	02	09	24	-34	14	00	-40	12	41	44	10	33	40	43	39	84		93			0.764**	
	10	20	07	35	-18	15	00	-17	05	41	27	-08	07	19	43	39	68		90			0.676**	
16	-03	02	09	22	-31	12	00	-36	11	37	40	09	30	36	39	35	76	93				0.846**	
	10	19	06	32	-17	14	00	-16	05	38	25	-07	07	17	39	36	63	94				0.735**	
17	-04	02	10	25	-36	14	00	-41	13	42	45	10	34	41	44	40						0.739**	
	10	20	06	35	-18	16	00	-18	05	41	27	-08	07	19	43	39						0.673**	
18	-03	02	09	23	-32	13	00	-38	12	38	41	09	31	38	40	37						0.810**	
	10	19	06	33	-17	15	00	-17	05	39	26	-07	07	18	40	37						0.719**	
19	-04	02	10	26	-37	15	00	-43	13	43	47	11	35	42	46	42	90					0.716**	
	12	24	07	41	-27	18	00	-20	06	48	32	-09	09	22	50	46	79					0.581**	
20	-04	03	14	35	-50	20	00	-58	18	59	64	14	48	58	62	57						0.527**	
	16	31	10	54	-28	24	00	-27	08	63	42	-12	11	29	66	60						0.442**	
Criterion	-03	02	07	19	-26	10	00	-31	10	31	34	08	25	30	33	30	64	79	71			Rural	
Correlation	07	14	04	24	-12	10	00	-12	03	28	18	-05	05	13	29	27	46	69	61			City	

Notes

1. The first two rows give the full model results.
2. For each model the upper row relates to the rural sample, the lower to the city sample.
3. Asterisks indicate a significant drop in R^2 compared with the full model, * = $P < 0.05$ ** = $P < 0.01$.
4. The three blocks within the table separate the biographic, dispositional and intellectual variables and models.
5. Full variable names are given in the method section.
6. The criterion correlation is the correlation between the predictor and the criterion, secondary reading achievement.

FIGURE 1
DETERMINANTS OF SECONDARY READING PERFORMANCE

Proportion of variance on achievement and progress attributable to full and reduced regression models

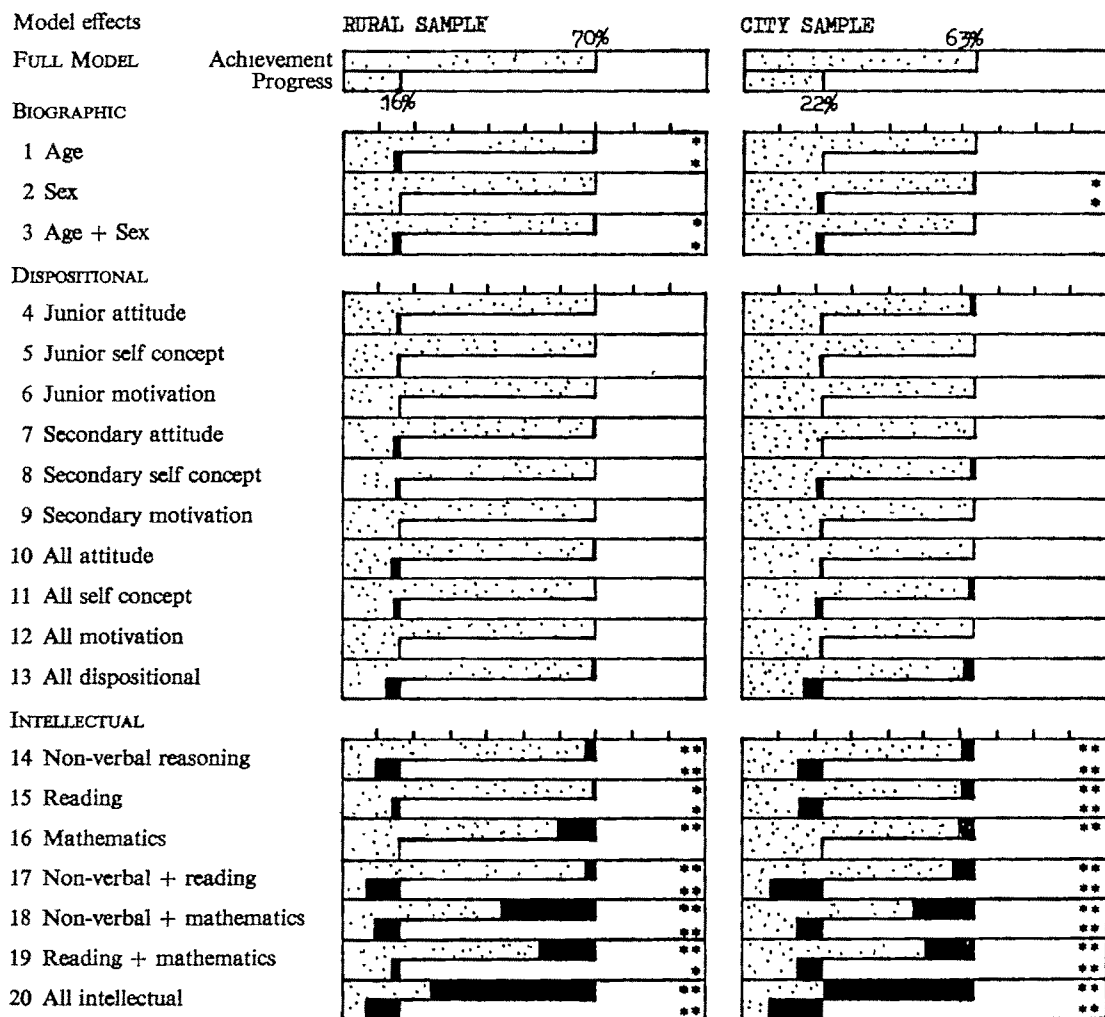


Notes

1. The full width of each horizontal bar represents 100 per cent of the criterion variance.
2. The speckled portion is the proportion attributable to each model.
3. The black portion is the difference between each model's variance and the full model's.
4. Significance markers (* $P < 0.05$, ** $P < 0.01$) give the F -test results testing the significance of the drop in the proportion of criterion variance predicted.
5. The top section of each bar represents achievement, the bottom section progress.

FIGURE 2
DETERMINANTS OF SECONDARY MATHEMATICS PERFORMANCE

Proportion of variance on achievement and progress attributable to full and reduced regression models



Notes

1. The full width of each horizontal bar represents 100 per cent of the criterion variance.
2. The speckled portion is the proportion attributable to each model.
3. The black portion is the difference between each model's variance and the full model's.
4. Significance markers (* $P < 0.05$, ** $P < 0.01$) give the F -test results testing the significance of the drop in the proportion of criterion variance predicted.
5. The top section of each bar represents achievement, the bottom section progress.

is implicit in the definition of progress since the residual score eliminates the effect of junior achievement. Otherwise the only effect that is not significant is for non-verbal reasoning with reading (Figure 1, city sample models 14 and 17).

The intellectual models investigate the effects of three predictors, non-verbal reasoning, prior achievement in the *same* subject and prior achievement in a *different* subject. For both achievement criteria prior achievement in the same subject has strongest effect. However when predictors are paired the emphasis changes. For reading the strongest pair is the two prior achievement measures, but for mathematics it comprises prior mathematics achievement and non-verbal reasoning. Comparability returns when all three predictors are used. In both instances there is a substantial increase in the criterion variance explained, the additional joint proportion exceeding 40 per cent for reading and mathematics. Results for the progress criteria are necessarily different because of the removal of prior achievement from the residual score. This really means that only two predictors are operative, prior achievement in the other subject and non-verbal reasoning. Again the full effect is relatively strong. Individual effects vary. For reading non-verbal reasoning is the stronger for the rural sample, mathematics for the city one. The mathematics progress effects match the achievement results with non-verbal reasoning being the more powerful. This pattern suggests that non-verbal reasoning is a more consistent determinant of progress than achievement in other subjects.

DISCUSSION

The uniformity of the findings for the two samples and for the two areas of performance makes extended discussion redundant. Probably the main point to stress is the assumption on which the analysis rests. This analysis sought to assess the nature of *independent* effects on early secondary school performance. Its conclusion that substantial independent effects are associated with intellectual measures, and conversely that non-intellectual measures (biographic or dispositional) do *not* independently determine performance, hinges on the covariance definition of independence. It implies that these measures effect performance even when other possible influences are controlled. Of course this conclusion must be limited to the domain of variables included here. It is not difficult to identify other non-intellectual variables (social class, introversion and stability are the most obvious candidates) whose effects may be strong enough to operate independently. Furthermore this conclusion does not imply that biographic or dispositional measures are not *related* to performance. Care was taken to point out the presence of many significant correlations for these relationships.

Principally the main conclusion of this study serves to remind that it is rarely satisfactory to isolate single determinants of success or failure during this period. It also confirms that apart from certain well-established factors such as prior achievement (Nisbet *et al.*, 1972) or intellectual ability (Dale and Griffith, 1965) most characteristics operate in association with others. This recalls the earlier finding (Youngman, 1978) which identified distinct and different profiles for pupils of similar ability. Indeed this network of interactions can be extremely complex. As the extended dispositional models show (Figures 1 and 2, model 13), even in combination attitude and personality effects may still be very weak. This corresponds with Behrens and Vernon's (1978) finding that personality variables had only low and inconsistent relationships with mathematics and reading achievement when multiple regression was used. Roach (1979) suggests that the effects of some cognitive variables may also disappear when other intellectual measures are controlled. Clearly if determinants other than intellectual ones are to be reliably identified, then extreme care is needed over the selection, measurement and analysis of pertinent characteristics.

REFERENCES

- BEHRENS, L. T., and VERNON, P. E. (1978). Personality correlates of over-achievement and under-achievement. *Br. J. educ. Psychol.*, 48, 290-297.
- COOLEY, W. W., and LOHNES, P. R. (1971). *Multivariate Data Analysis*. New York: John Wiley.
- DALE, R. R., and GRIFFITH, S. (1975). *Down Stream: A Study of Failure in the Grammar School*. London: Routledge, Kegan and Paul.
- NEAL, P. D. (1975). *Continuity in Education: EDC Project five*. Educational Development Centre, City of Birmingham Education Department.
- NISBET, J. D., and ENTWISTLE, N. J. (1966). *The Age of Transfer to Secondary Education*. London: University of London Press.
- NISBET, J. D., and ENTWISTLE, N. J. (1969). *The Transition to Secondary Education*. London: University of London Press.
- NISBET, J. D., WELSH, J., and ENTWISTLE, N. J. (1972). Age of transfer to secondary education: a postscript. *Br. J. educ. Psychol.*, 42, 233-239.
- PUMFREY, P. D., and WARD, J. (1976). Adjustment from primary to secondary school. *Educ. Res.*, 19, 25-34.
- ROACH, D. A. (1979). The effect of conceptual style preference, related cognitive variables and sex on achievement in mathematics. *Br. J. educ. Psychol.*, 49, 79-82.
- SUMNER, R., and BRADLEY, K. (1977). *Assessment for Transition: A Study of New Procedures*. Windsor: NFER.
- THOMPSON, B. (1975). Adjustment to school. *Educ. Res.*, 17, 128-136.
- VELDMAN, D. J. (1967). *FORTRAN Programming for the Behavioral Sciences*. New York: Holt, Rinehart and Winston.
- YOUNGMAN, M. B. (1976). *Programmed Methods for Multivariate Data*. Version Five. School of Education, University of Nottingham.
- YOUNGMAN, M. B. (1978). Six reactions to school transfer. *Br. J. educ. Psychol.*, 48, 280-289.
- YOUNGMAN, M. B. (1979). *Analysing Social and Educational Research Data*. Maidenhead: McGraw-Hill.
- YOUNGMAN, M. B., and LUNZER, E. A. (1977). *Adjustment to Secondary Schooling*. School of Education, University of Nottingham.

(Manuscript received 17th September, 1979).

THE RELATIONSHIP BETWEEN LOCUS OF CONTROL AND ACADEMIC ACHIEVEMENT, ANXIETY, AND LEVEL OF ASPIRATION

BY D. BAR-TAL, D. KFIR, Y. BAR-ZOHAR AND M. CHEN

(*School of Education, Tel-Aviv University*)

SUMMARY. While holding the influence of socio-economic status constant, the relationship between locus of control and academic achievement, anxiety, and level of aspiration was examined among 2,438 ninth grade Israel-Jewish students of Asian or African background and European, American or Israeli backgrounds. It was observed that, in general, internals tend to attain greater academic achievement, to express less anxiety, and to have higher level of aspiration. The relationship between academic achievement, level of aspiration, and the perception of locus of control appears to be accentuated among students of Asian or African origin to a greater extent than among students of European, American, or Israeli origin.

INTRODUCTION

RESEARCH in the area of locus of control has largely focused upon the investigation of the relationship between the perception of locus of control and various behaviours, attitudes, and/or personality characteristics (for comprehensive reviews see Lefcourt, 1966; Rotter, 1966; Joe, 1971; Phares, 1975). The present study investigated the relationship between locus of control, on the one hand, and academic achievement, anxiety, and level of aspiration, on the other hand, while controlling the influence of socio-economic status. It explored the aforementioned relationships in a sample of the total population of ninth grade junior high school Israeli-Jewish adolescents within a cross-cultural context. The present study compared Israeli children from two different subcultures and it facilitates a comparison with results found in research performed in the USA.

Since the publication of the Coleman report (Coleman *et al.*, 1966) which indicated that, at least with respect to minority students, locus of control is firmly related to academic achievement, more than 40 studies have investigated the relationship between perception of locus of control and academic achievement (see Bar-Tal and Bar-Zohar, 1977, for an extensive review). Although the results are somewhat inconsistent, internal perception of control tends to be positively correlated with academic achievement. On the basis of these findings it has been suggested that perception of locus of control is an important personality variable which predicts academic success in schools (e.g., Morrison and McIntyre, 1971; Phares, 1975). This view is strengthened by the evidence that internal perception of control is systematically related to behaviours which increase the probability of successful academic performance (cf. Bar-Tal and Bar-Zohar, 1977).

Locus of control also appears to be associated with anxiety. More specifically, numerous investigators have reported that anxiety is positively related with an external orientation (e.g., Watson, 1967; Allen *et al.*, 1974). Conceivably, perception of lack of control over one's environment induces a certain amount of anxiety. This anxiety may be bolstered by the external's tendency to perceive external controlling forces as malevolent and untrustworthy to a greater extent than the internal's tendency (e.g., Lamont, 1972; Levenson and Mahler, 1975).

In addition, perception of locus of control has been observed to be related to level of aspiration. Gurin *et al.* (1969) and Lao (1970), for example, observed that students whose orientation was internal were characterised by higher educational expectations and aspirations than students whose orientation was external. It is possible that

individuals who feel that they can control their own environment also believe that, since success depends on them, they can achieve more than externals who feel that they are dependent on the whim and desires of other people or events.

Numerous studies have focused on child rearing patterns which appear to influence the perception of locus of control. In this vein, internal children were found to have warmer, more protective, more consistent and more realistic parents than external children (e.g., Baumrind, 1966; Katokvsky *et al.*, 1967; Nowicki and Segal, 1974; Loeb, 1975). Socio-economic classes tend to differ in their child rearing patterns (e.g., Sears *et al.*, 1957). It follows, accordingly, that children from different socio-economic levels may differ in their perceptions of locus of control. Indeed, a number of studies have pointed to a strong relationship between perception of locus of control and socio-economic status. In general, the evidence shows that individuals of low socio-economic status tend to be more external than individuals of high socio-economic status (e.g., Battle and Rotter, 1963; Gore and Rotter, 1963; Shaw and Uhl, 1971).

Given that locus of control as well as academic achievement appear to be closely related to socio-economic status (e.g., Coleman *et al.*, 1966), one may claim that the positive relationship between perception of locus of control and academic achievement is spurious. That is, the relationship between locus of control and academic achievement may be a result of their common relationship with socio-economic status.

The main purpose of the present study was to examine the relationship between locus of control and academic achievement, anxiety, and level of aspiration, while holding the influence of socio-economic status constant. The above described investigation was made within the framework of a longitudinal survey carried out during 1971-1975 (Chen *et al.*, 1978). This survey intended to evaluate the effect of the educational reform being implemented in Israel, which changed the schooling system from eight years of elementary education and four years of secondary education to six years of elementary education, three years of junior high school education, and three years of senior high school education. The main purpose of the reform was to raise the academic achievement of the students of Asian-African origin and to increase their social integration by establishing regional-integrative junior high schools. This study compared two ethnic groups of Israeli pupils. The ethnic origin of the pupils was defined dichotomously in accordance with their fathers' extraction. More specifically, students of Asian or African origin (AA) included all children whose fathers were of Asian or African extraction, excluding fathers born in Israel, whereas students of European, American, or Israeli origin (EAI) included all children whose fathers were of European or American extraction, or whose fathers were born in Israel. These ethnic classifications are commonly used with reference to the Jewish population by the Israeli Census Bureau Abstracts and in research studies.

In comparison with the Western community, the Eastern community is clearly in a disadvantageous position. The Eastern community tends to belong to a lower socio-economic level, has fewer years of formal education, and tends to be represented to a lesser extent in the more influential national political institutions. Within this relatively dichotomous framework, social researchers have reported psychological reactions that differentiate AA from EAI children (cf., Peres, 1976; Lewy and Chen, 1977; Minkowitz *et al.*, 1977). Thus, the second purpose of the present study is to compare the relationships among the investigated variables between the two ethnic groups, while controlling for the effect of socio-economic status.

In sum, the following hypotheses were tested: (a) while socio-economic status is held constant, locus of control will be related to academic achievement, anxiety, and level of aspiration in a direction that indicates that the more internal locus of control will be associated with higher academic achievement, lower level of anxiety, and higher

level of aspiration; (b) the above described relationship between locus of control and academic achievement will be especially accentuated among AA students. The latter prediction was derived from studies which indicate that locus of control is a better predictor of academic success among certain ethnic minorities in the USA, whose SES tends to be low, than among white children of either high or low SES (e.g., Coleman *et al.*, 1966).

METHOD

Subjects

Subjects were 2,438 ninth grade students who participated in the longitudinal study of Israeli Jewish high school students (Chen *et al.*, 1978). These students studied in 104 ninth grade classes of 19 junior high schools. Fifty per cent (1,216) of the subjects were of AA origin. The subjects constituted a stratified random sample of the total population of ninth grade Jewish students in Israel. However, since some students did not answer all the questions, the actual number answering each question is smaller than the total number. Therefore, in the analyses which required more than one variable, only the data of those students who answered all the questions were used.

Instruments

The instrument which assessed various attitudes and beliefs toward the school, family, society, education and self was a closed-ended questionnaire. Its items were grouped as variables on the basis of a factor analysis and content analysis (for details of these analyses and the procedures see Chen *et al.*, 1978). The following variables were used in the present study.

Variables

Locus of control. Six items were designated to measure locus of control. These items were adapted from Coleman *et al.* (1966), Rotter (1966) and Gurin *et al.* (1969). The index which was obtained (Cronbach's $\alpha = 0.53$) enabled us to rank individuals on a seven-point scale ranging from 0 to 6. A low score on the scale indicated that the individual tends to be external, while a high score meant that the individual tends to be internal. The six items were: (a) the life of a person is pre-determined and he does not have the power to change it; (b) achievement of a desirable goal depends mainly on being in the right place at the right time; (c) I find that whatever was determined by destiny, actually happens; (d) many times questions on an examination tend to be so unrelated to course work that studying is really useless; (e) many times I feel that I do not have control over the direction my life is taking; (f) luck is no less important in success than hard work.

Socio-economic status. The index of socio-economic status (Cronbach's $\alpha = 0.54$) was assessed by a scale ranging from 3 to 20. A high score on the scale indicated that a student has high SES, and consisted of three items which were determined using factor analysis. The three items asked for the number of siblings, father's level of education, and number of books at home. (Number of children in a family is used as an indicator of socio-economic status by the Israeli Census Bureau.)

Academic achievement. Academic achievement was measured by a battery of four standardised achievement tests in Mathematics, English, Reading Comprehension, and Bible. Percentage of items answered correctly on all four tests constituted the achievement score.

Manifest school anxiety. Eight items adapted from Sarason's Test Anxiety Scale for Children (1960) were used to measure school anxiety (Cronbach's $\alpha = 0.68$), e.g., "When the teacher says that she is going to call upon some boys and girls in the class to do arithmetic problems, do you hope that she will call upon someone

else and not you?", or "When you are taking a test, does the hand you write with shake a little?". The scores ranged from 8 to 16, and the low score indicated high anxiety.

Level of aspiration. Four questions measured level of aspiration (Cronbach's $\alpha = 0.56$): (a) "Will you try to take the matriculation examination?";* (b) "Do you believe that you will succeed in the matriculation examination?". In both of the above questions, each subject was asked to reply either "Yes" or "No". The third question was: "When you finish junior high school,† where do you plan to study?". The fourth question inquired about the future occupation or profession desired by the student. The scores ranged from 4 to 17 and the high score indicated high level of aspiration.

RESULTS

Correlations between locus of control and each of the other variables were calculated separately for AA and EAI students (see Table 1). The results indicate that locus of control is related to SES, academic achievement, school anxiety, and level of aspiration among AA and EAI students. That is, students who tended to be internal have higher SES, higher academic achievement, lower level of anxiety and higher level of aspiration. Comparisons between the correlations of the AA and EAI students by *t*-tests for independent correlations, reveal that only the correlation between academic achievement and locus of control is significantly higher in the case of AA students than with respect to EAI students ($t(2263) = 3.25, P < 0.001$).

TABLE 1
CORRELATIONS OF LOCUS OF CONTROL WITH SES ACADEMIC ACHIEVEMENT, MANIFEST SCHOOL ANXIETY, AND LEVEL OF ASPIRATION FOR STUDENTS OF ASIAN OR AFRICAN ORIGIN AND STUDENTS OF EUROPEAN, AMERICAN AND ISRAELI ORIGIN

Variables	Locus of Control	
	Asian-African Students (N=1096)	European-American-Israeli Students (N=1169)
S.E.S.	0.23*	0.21*
Academic achievement	0.41*	0.29*
Manifest school anxiety	0.14*	0.19*
Level of aspiration	0.32*	0.26*

* $P < 0.001$

Note. The direction of scores of each variable is described in the Method section.

In addition, partial correlations between locus of control and other variables were calculated in order to partition the influence of socio-economic status (see Table 2). Table 2 shows that although the magnitude of the correlations was somewhat reduced after the removal of the effect exerted by SES, the relationships still remain significant. Especially high correlations of locus of control with academic achievement and level of aspiration for AA were observed. Correlations between academic achievement and locus of control as well as between level of aspirations and locus of control were

* Matriculation is a comprehensive examination in five subject areas usually administered upon completion of high school. This examination is a requirement for entrance to Israeli colleges and universities.

† At the time of this study, Israeli compulsory education terminated, with few exceptions, at the end of junior high school. Therefore, after the ninth grade most of the students then chose either to continue their studies in a vocational, or academic programme, or to discontinue their schooling.

TABLE 2

PARTIAL CORRELATION OF LOCUS OF CONTROL WITH ACADEMIC ACHIEVEMENT, MANIFEST SCHOOL ANXIETY, AND LEVEL OF ASPIRATION CONTROLLING FOR SES AMONG ASIAN-AFRICAN AND EUROPEAN AMERICAN-ISRAELI STUDENTS

Variables	Locus of Control	
	Asian-African Students (N=1096)	European-American-Israeli Students (N=1169)
Academic achievement	0.35*	0.22*
Manifest school anxiety	0.11*	0.18*
Level of aspiration	0.27*	0.19*

* $P < 0.001$

significantly higher in the case of AA students than EAI students ($t(2263) = 3.37$, $P < 0.001$, and $t(2263) = 1.99$, $P < 0.05$, respectively).

A multiple regression technique was used in order to examine the relative contribution of perception of locus of control to academic achievement, anxiety, and level of aspiration, while controlling for socio-economic status.

Table 3 shows that percentage of explained variance by socio-economic status is significant for all the variables except for school anxiety in the case of AA subjects. The socio-economic status variable is especially powerful in accounting for the variance of academic achievement in the two ethnic groups, and in explaining level of aspiration for the AA Group. Referring to Table 3, perception of locus of control, a personality

TABLE 3

SUMMARY OF MULTIPLE REGRESSION ANALYSES

Dependent Variables	Independent Variables			
	First Step: Socio-economic Status		Second Step: Locus of Control	
	Percentage of Explained Variance	Standardised Beta Weight	Percentage of Explained Variance	Standardised Beta Weight
<i>Asian-African Group</i>				
Academic achievement	17.1**	-0.34**	9.7**	0.32**
Anxiety	0.2	-0.00	1.4**	0.12**
Level of aspiration	8.8**	0.23**	7.2**	-0.28**
<i>European-American-Israeli Group</i>				
Academic achievement	26.0**	-0.47**	3.7**	0.20**
Anxiety	0.4	-0.03	2.3**	0.16**
Level of aspiration	19.9**	0.41**	2.7**	-0.17**

* $P < 0.05$ ** $P < 0.01$

Note. The direction of the scores of each variable is described in Method section.

variable, seems to be related to academic achievement, anxiety, and level of aspiration for both ethnic groups. The percentage of explained variance by locus of control is significant for all the dependent variables. In addition, it appears that standardised beta weights of school anxiety are greater for perception of locus of control than for socio-economic status. Moreover, it appears that the relative weight of perception of

TABLE 4

MEANS, STANDARD DEVIATIONS AND PAIRWISE NON-ORTHOGONAL COMPARISONS BETWEEN STUDENTS OF ASIAN ORIGIN AND EUROPEAN-AMERICAN-ISRAELI ORIGIN

Variables	Asian-African Students			European-American Israeli Students			<i>t</i>
	N	M	SD	N	M	SD	
Locus of control	1071	9.37	1.51	1155	10.06	1.31	11.65*
Socio-economic status	1078	10.45	3.09	1132	14.84	2.84	43.45*
Academic achievement	1158	49.75	17.11	1145	65.26	14.99	23.82*
Manifest school anxiety	1086	13.86	1.96	1148	13.96	1.84	1.32
Level of aspiration	1086	12.54	2.79	1157	14.37	2.14	17.47*

* $P < 0.001$

Note. The direction of the scores of each variable is described in the Method section.

locus of control in explaining the variance of academic achievement and level of aspiration is greater for the AA group than for the EAI group.

Finally, AA and EAI students were compared on all the variables (see Table 4). The results indicate that AA students are more external, have lower SES, have lower academic achievement, and have lower level of aspiration than EAI students. A difference between the two groups in the level of school anxiety was not detected.

DISCUSSION

The results of the present study confirmed the hypotheses concerning the differences between internal and external students. Internal students were found to have higher academic achievement, to have less anxiety, and higher level of aspiration than external students. These results are of especial interest in view of the analyses which controlled the effect of socio-economic status. Thus, reaching beyond the influence of socio-economic status, perception of locus of control was found to be a significant predictor of academic achievement, manifest/school anxiety, and level of aspiration. It seems that the above psychological differences between internal and external children can, at least partially, contribute to our understanding of external children's inferiority with respect to academic achievement. High anxiety and low level of aspiration are generally viewed as dysfunctional to academic performance. External students who feel that they cannot change events, tend to be worrisome, anxious, and to have low expectations for future success. Thus, perception of locus of control appears to be a powerful variable related to the above-mentioned psychological variables and academic achievement even when socio-economic status is controlled. These results are in line with findings which indicate that sense of control over the environment is related to academic achievement (e.g., Coleman *et al.*, 1966; Gurin *et al.*, 1969; Lao, 1970).

The present study replicates results obtained in the USA which suggest that perception of locus of control is an especially meaningful variable among minority groups. Coleman *et al.* (1966) findings indicated that with respect to students from minority groups the perception of locus of control is more strongly related to academic achievement than any other psychological variable. In Israel, similar results were obtained: the relationship between the perception of control with achievement and level of aspiration appears to be accentuated among AA students to a greater extent than among EAI students. One possible explanation of the above result is based on the assumption that the impact of locus of control diminishes as other elements that are conducive to academic success and level of aspiration increase in strength. Among EAI students, most of whom are middle class, achievement and level of aspiration are

supported by numerous variables, such as history of reinforcement, availability of models, positive self-image, etc. Consequently, there is a diminishing function of the perception of locus of control as a determinant of academic achievement and level of aspiration. Among AA students, however, there are fewer variables available to support academic achievement and level of aspiration. As a result, the role of perception of locus of control becomes more dominant and critical as a determinant of success in school.

Finally, the results of the present study indicate that AA students are more external, have lower SES, have lower academic achievement and lower level of aspiration than EAI students. With regard to locus of control, it should be pointed out that the externality of the AA students could originate either from the 'individual' or 'system' blame. In the first case the person blames his own destiny or luck and in the latter case he blames the social system for his fate. Gurin *et al.* (1969) reported that black college students with individual blame performed in a superior fashion on achievement tests, collegiate courses, and programme tasks than students whose external orientations were ideological. Future studies should make the described distinction while investigating the locus of control of the AA students in Israel.

[The study reported above is a part of the research about Israeli junior high schools financed by the Israeli Ministry of Education. Requests for reprints should be sent to Daniel Bar-Tal, School of Education, Tel-Aviv University, Tel-Aviv, Israel.]

REFERENCES

- ALLEN, G. J., GIAT, L., and CHERNEY, R. J. (1974). Locus of control, test anxiety, and student performance in a personalised instruction course. *J. educ. Psychol.*, 66, 968-973.
- BAR-TAL, D. and BAR-ZOHAR, Y. (1977). The relationship between perception of locus of control and academic achievement: review and some educational implications. *Contemporary educ. Psychol.*, 2, 181-199.
- BATTLE, E., and ROTTER, J. B. (1963). Children's feelings of personal control as related to social class and ethnic group. *J. of Personality*, 31, 482-490.
- BAUMRIND, D. (1966). Effects of authoritative parental control on child behavior. *Child Dev.*, 37, 887-907.
- CHEN, M., LEWY, A., and ADLER, C. (1978). *Process and Outcome of Education: Evaluation of the Contribution of the Junior High Schools in the Educational System*. Report to Israeli Ministry of Education. (Hebrew.)
- COLEMAN, J. S., CAMPBELL, E. Q., HOBSON, C. J., MCPARTLAND, J., MOOD, A. M., WEINFELD, F. D., and TORK, R. L. (1966). *Equality of Educational Opportunity*. Superintendent of Documents Catalog No. FS 5238: 38001, US Government Printing Office, Washington, D.C.
- GORE, P. M., and ROTTER, J. B. (1963). A personality correlate of social action. *J. Personality*, 31, 58-64.
- GURIN, P., GURIN, G., LAO, R. C., and BEATTIE, M. (1969). Internal-external control in the motivational dynamics of Negro youth. *J. social Issues*, 25, 29-53.
- JOE, V. C. (1971). Review of the internal-external control construct as a personality variable. *Psychol. Rep.*, 28, 619-640.
- KATKOVSKY, W., CRANDALL, V. C., and GOOD, S. (1967). Parental antecedents of children's beliefs in internal-external control of reinforcements in intellectual achievement situations. *Child Dev.*, 38, 766-776.
- LAMONT, J. (1972). Depression, locus of control and mood response set. *J. clin. Psychol.*, 28, 342-345.
- LAO, R. C. (1970). Internal-external control and competent and innovative behavior among Negro college students. *J. Person. soc. Psychol.*, 14, 263-270.
- LEFCOURT, H. M. (1966). Internal versus external control of reinforcement: a review. *Psychol. Bull.*, 65, 206-220.
- LOEB, R. C. (1975). Concomitants of boys' locus of control examined in parent-child interaction. *Dev. Psychol.* 11, 353-358.
- LEVENSON, H., and MAHLER, I. (1975). Attitudes toward others and components of internal-external locus of control. *Psychol. Rep.*, 36, 209-210.
- LEWY, A., and CHEN, M. (1977). Differences in achievement: a comparison over time of ethnic group achievement in the Israeli elementary school. In CHOPPIN, B. H., and POSTLETHWAITE, T. N. (Eds.), *Evaluation in Education: International Progress*. Oxford: Pergamon Press.
- MINKOWITZ, A., DAVIS, D., and BASBI, J. *Evaluation of Educational Achievements in Elementary Schools in Israel*. Report to Israeli Ministry of Education. (Hebrew.)

- MORRISON, A., and MCINTYRE, D. (1971). *Schools and Socialization*. Baltimore: Penguin.
- NOWICKI, S., and SEGAL, W. (1974). Perceived parental characteristics, locus of control orientation, and behavioral correlates of locus of control. *Dev. Psychol.*, 10, 33-37.
- PERES, Y. (1976). *Ethnic Relations in Israel*. Tel Aviv: Sifriat Hapoalim. (Hebrew.)
- PHARES, E. J. (1975). *Locus of Control in Personality*. Morristown, N.J.: General Learning Press.
- ROTTER, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monogr.*, 80 (1, Whole No. 609).
- SARASON, S. B., DAVIDSON, K. S., LIGHTHALL, F. F., WAITE, R. R., and RUEBUSH, B. K. (1960). *Anxiety in Elementary School Children*. New York: John Wiley & Sons.
- SEARS, R. R., MACCOBY, E. E., and LEVIN, H. (1957). *Patterns of Child Rearing*. Evanston, Ill.: Row, Peterson.
- SHAW, R. L., and UHL, N. P. (1971). Control of reinforcement and academic achievement. *J. educ. Res.*, 64, 226-228.
- WATSON, D. (1967). Relationship between locus of control and anxiety. *J. Person. soc. Psychol.*, 6, 91-92.

(Manuscript received 8th March, 1979)

The British Journal of Psychology

Volume 70, Part 4, December 1979

Selected contents

- M. Nagpel & B. S. Gupta. Personality reinforcement and verbal operant conditioning
- K. I. Manktelow & J. St B. T. Evans. Facilitation of reasoning by realism
- W. R. Crozier. The interaction of value and subjective probability in risky decision-making
- A. C. Owens & J. T. E. Richardson. Mental imagery and pictorial memory
- M. G. Eley. The suitability of placement tasks as analogues for syllogistic reasoning
- B. J. Fellows. The British use of the Barber Suggestibility Scale
- A. Philpott & J. Wilding. Semantic interference from subliminal stimuli in a dichoptic viewing situation
- K. Millar. Noise and the 'rehearsal-masking hypothesis'
- Book reviews

Volume 71 (1980) £35.00 (US\$77.00). Single parts cost £11.00 (US\$24.00)

Volume 70 (1979) £26.00

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN, UK

VISUAL AND OCULAR CONTROL ANOMALIES IN RELATION TO READING DIFFICULTY*

By C. H. BEDWELL

(*The City University, London*),

R. GRANT

(*West Sussex LEA*)

AND J. R. MCKEOWN

(*East Sussex LEA*)

SUMMARY. The visual behaviour under both static and dynamic viewing conditions was examined in a group of 13-year-old successful readers, compared with a group of the same age retarded in reading. Each group was video recorded reading aloud, without sound, a passage on the Neale Analysis that was well within their capacity to read. Criteria for selection and experimental controls are described. Each group was assessed for letter chart visual acuity, and examined on the full range of tests on a standard vision screener. The range of fusion that could be obtained before binocular vision broke down was assessed using polarised vectographs. Each pupil was asked whether vision tended to become double when reading. Dynamic visual behaviour, ocular control, and associated postural compensation and muscular stress while actually reading were investigated. The data obtained from these static and dynamic tests were then compared for the successful and retarded reading groups. It was found that the group retarded in reading indicated a significant relationship between anomalies of dynamic binocular viewing and their reading attainment. On the tests under static viewing only one significant difference—poor stereopsis—could be found between the two groups. Research supports the notion that problems of dynamic binocular vision and control while reading are important.

INTRODUCTION

THERE is disagreement in the literature over the relative importance and significance of visual processes in reading attainment. The area of controversy relates to the relative importance of function at the level of the eye, as the periphery of the visual system, binocular behaviour, and eye-movement control, compared with the visual processing occurring at the higher levels of brain functioning. The issue has been described by Rayner and McConkie (1977) as the direct-perception view, in contrast to hypothesis testing in the form of psycho-linguistic processing using contextual clues and predictions (Goodman, 1975).

In one of the first papers on the visual aspects of reading difficulty, Hinshelwood (1917) concluded that minor refractive errors and imbalances of the extra-ocular muscles were not significant in causing reading and writing problems. In the case of the child with severe reading disability Keeney (1968) and his co-authors also felt that visual acuity, refractive error and minor binocular imbalances were not significant factors. But the binocular variables included in this study were perhaps too simple.

If a somewhat more complex view on binocular disfunction is considered in relation to reading difficulty the evidence provided by different workers suggests that binocular factors might be of greater importance. In a study on reading performance viewing monocularly and binocularly, Birnbaum (1968) found that there was a significant number of children with reading disability with a reduced performance on binocular viewing that was related to binocular efficiency. Spache (1976) reported that binocular disfunction appeared to be significant in relation to reading difficulty, and

* Copies of this paper are obtainable from C. H. Bedwell, 33a Station Road, New Barnet, Herts. EN5 1PH. (S.A.E., please)

recommended that a binocular reading test should be incorporated as part of the investigation. In contrast Fletcher (1969) felt that there was little evidence that lack of single binocular vision, or the presence of strabismus, amblyopia, or poor retinal rivalry, were causes of learning disability, though disorders of visual perception appeared to affect ability.

Some researchers have looked into the effects of binocular training on reading. Taylor (1966) demonstrated how eye-movement and range of fusion affected reading difficulty, and suggested that reading skill could be improved by appropriate visual management. Hirsch and Anderson (1976), however, in a summary article in this field, concluded that the general use of visual motor training has inadequate experimental control, and results should therefore be treated as inconclusive.

Numerous other articles have examined related aspects. Flax (1968) concluded that end-organ visual functions, such as fusion, convergence, and accommodation, can contribute to reading difficulty, but are more important in maintenance of efficiency of reading. Visual form perception, visual memory, appreciation of visual direction differences, visuo-motor control, and eye-hand co-ordination appear to play a role in reading difficulties. Goldberg and Guthrie (1972) also found that a significant positive correlation existed between visual perception, visual memory, and visual sequencing, in relation to reading difficulty.

Much greater emphasis has been put by some research conclusions on higher processing. Critchley (1964) felt that specific developmental dyslexia was a genetically determined constitutional disorder and was not visual in origin, and this view has been supported by other workers (e.g., Goldberg *et al.*, 1970). Buchanan (in Keeney and Keeney, 1968) reviewed the evidence for cortical localisation for dyslexia, and concluded that as yet there is no objective means by which an anatomical or physiological lesion could be displayed as the underlying cause.

A discussion of the significance of cerebral hemispherical dominance and the effects of cross-laterality are outside the scope of this paper except as they might refer to ocular dominance. The type of tests (e.g., sighting) commonly employed to indicate ocular dominance in cross-laterality studies tend to give inconsistent findings in relation to reading difficulty. The situation is complicated because of the manner in which the retina is represented in the cerebral hemispheres, and that both cerebral hemispheres are likely to be stimulated in most viewing situations. In a controlled experiment, however, Wheeler *et al.* (1977) found that a cross-lateral group performed significantly less well than a group exhibiting normal laterality in identifying units of information presented simultaneously in the two halves of the field.

The role and significance of binocular functions is a difficult area to investigate and to understand. Earlier experience by Bedwell (1973, 1978) in examining children with reading difficulty had indicated that complex dynamic binocular visual aspects are likely to affect many aspects of reading difficulty. These research projects were with children in infant and junior schools, and it was felt that this research should now be extended into middle schools.

METHOD OF VISUAL INVESTIGATION AND RESULTS

This paper describes an investigation into the relationship between binocular function, ocular control and reading difficulty in 40 children with a mean age of 13 : 7 years.

Tests of binocular function

Initially vision was examined using a Snellen visual acuity letter chart and then a Titmus vision screener, i.e., both tests involving static viewing. Such methods would commonly be used to investigate vision during a school medical examination. These

tests were then supplemented by an assessment of the range of fusion, that is the degree of out-of-alignment the eyes could contain while looking at an object and yet maintain single binocular vision. A record was also made showing which pupils experienced letters or words going double while actually reading. Finally dynamic visual behaviour, binocular eye control (together with any associated compensatory head movements) and ocular stress, were then investigated while the pupil was actually looking between different targets, and while reading. By this means it was possible to assess the significance of visual examination both under static and dynamic conditions.

Sample

Forty children aged between 13 and 14 years attending one large secondary school were tested to establish reading age and spelling age. Twenty-five children were found to be two or more years retarded in reading, and are described as poor readers. The other pupils were reading at their age level and above, and are described as good readers. The criterion for selection of both groups was that they should all have singular dominance, either left or right. Both groups approximated to a national average of 10 per cent singular left dominance (Naidoo, 1972). Reading age was assessed using the Schonell R4 Test (Schonell, 1950), and spelling by Schonell dictation Test E (Schonell, 1950). The characteristics of the two groups are given in Table 1.

TABLE 1
READING AND SPELLING AGES

Average Group Characteristics	Poor Readers (N = 25)		Good Readers (N = 15)	
	Mean	(SD)	Mean	(SD)
Chronological Age in Years	13.7	(0.38)	13.7	(0.26)
Reading Age Schonell R4	10.7	(0.44)	13.6	(0.77)
Spelling Age Schonell Dictation Test E	9.8	(1.1)	13.8	(0.77)

Experimental controls

Each individual was placed in random order for the purposes of administering and assessing the tests. The experimental procedures for static and dynamic viewing situations are outlined below. Each group of visual tests was analysed independently of the other. Random order and no sound ensured that actual reading ability did not influence interpretation of the data.

VISUAL TESTS AND RESULTS OBTAINED

Letter-chart visual acuity

The monocular and binocular visual acuities were assessed at the conventional distance viewing of 6 m on a standard Snellen visual acuity chart. This method of vision testing is that which is normally used in a conventional ophthalmic examination, and as part of a school medical examination. A visual acuity of 6/9 or less was taken as the point at which the pupil had failed to pass the test.

The data obtained and indicated in Table 2 show that static distance visual acuity does not appear to be significant in relation to reading difficulty (χ^2 test).

Vision screening

Pupils were examined on the standard range of vision tests incorporated in the Titmus Vision Screener. In some areas of the country a vision screener is used to assess vision as part of a school medical examination instead of the classical Snellen

TABLE 2
RELATIONSHIP BETWEEN VISUAL ACUITY AND READING AGE

	Passed	6/9 or less Failed	Total
Poor Readers	19	6	25
Good Readers	14	1	15
Total	33	7	40

visual acuity chart. The Titmus instrument used for this project is one of a number of screeners commonly used in the United States and Canada, and to a lesser extent in this country, as a means of examining vision. In most vision screeners visual acuity is assessed binocularly and monocularly using a non-literate symbol—in the case of the Titmus instrument a Landolt C. The subject views diamond patterns of one O and three Cs. As they become smaller he has to indicate in which position there is a complete circle in one of the four positions forming a diamond. For distance vision the targets are placed optically at infinity, and for near vision so that they appear 33 cm from the eye, therefore requiring approximately the degree of accommodation or focusing, and also convergence, that would be used when reading or writing.

There are also dissociation tests which indicate whether the eyes tend to drift out of line horizontally for distance and near fixation, and also vertically at distance. They are called dissociation tests as the targets seen, e.g., an arrow by one eye and a scale by the other, do not have to be fused as one superimposed image.

Incorporated also are tests to indicate suppression of vision in one eye. Stereoscopic visual acuity is assessed by noting which of a series of circles is seen in space in front of the others.

In many cases involving the use of a vision screener for vision testing of school children a simpler modified vision test is used. As 13-year-old pupils were being examined it was felt that it would be desirable to use the normal full adult range of tests.

The data obtained on visual acuity for distance and near vision using a vision screener did not support any association between reading ability any more than did letter chart visual acuity.

The results shown in Table 3 demonstrate that any tendency for the eyes to deviate horizontally on static viewing at distance or near to, as recorded on a dissociation type of test for examining binocularity, does not appear to be significant in relation to reading difficulty. There does, however, appear to be a relationship between poor stereoscopic visual acuity and reading difficulty.

Range of fusion

Range of fusion and any tendency for suppressing the images in one or other eye are aspects that it is usual to examine when investigating binocular function more deeply.

The range of fusion, or the degree to which the eyes can drift out of line and still see a single image, was assessed using polarised vectograms. This test comprises two photographs that are superimposed, and viewed through polarised goggles. A normal person can see the scene in depth, or stereoscopically. The polarisation effect is so produced that with one eye a picture can be seen on one photograph, but the other appears transparent. The other eye at the same time sees a picture on the second photograph, but the first appears transparent. The ability of the eyes to diverge

TABLE 3

RELATIONSHIP BETWEEN READING AGE AND VARIABLES DERIVED FROM THE TITMUS SCREENER

Variable	Good Readers N = 15		Poor Readers N = 25		t value	P < (One tailed)
	Mean	(SD)	Mean	(SD)		
Vertical imbalance for distance viewing	5.0	(0.0)	4.32	(1.55)	0.23	NS
Horizontal imbalance for distance viewing	10.13	(0.92)	10.25	(1.18)	0.09	NS
Horizontal imbalance for near viewing	8.00	(2.48)	8.33	(2.84)	1.17	NS
Stereovision	8.73	(2.25)	7.00	(3.11)	2.87	0.05

(denoted by letters) or to converge (denoted by numbers) relative to one another, and still see a single scene in depth, can be assessed by moving the vectograms out or in until stereoscopy is lost and the images go double.

Any tendency for one of the polarised vectograph images to appear dimmer, or not to be seen, would indicate a tendency for suppression of vision in the weaker eye.

Statistical analysis, shown in Table 4, on the significance of these aspects alone of fusion in relation to reading difficulty, indicated that neither suppression, the ability to

TABLE 4

RELATIONSHIP BETWEEN FUSIONAL ABILITY AND READING AGE

Polarised Vectograms	Percentage Numbers Showing Abnormalities		χ^2	P < (One tailed)
	Good Readers	Poor Readers		
Dimming or loss of image seen by one eye (suppression)	80	56	2.373	NS
Ability to converge before fusion breaks (numbers)	46.6	58	1.777	NS
Ability to diverge before fusion breaks (letters)	16.0	33	1.615	NS

maintain fusion on convergence, or on divergence, were significant in relation to reading difficulty.

Double vision

If while reading, letters or words become double, this is an indication that single binocular vision has broken down. Usually the possibility of such a phenomenon

occurring may not have been realised by the teacher, and as the pupil may always have had such a tendency, the incidence and significance of double vision while reading may have gone unnoticed.

Double vision while reading is one of a number of manifestations of binocular embarrassment. Table 5 presents details of its reported presence among the pupils in each group. Double vision is one of the readily identifiable problems that the teacher might notice as being relevant to reading difficulty.

TABLE 5
RELATIONSHIP BETWEEN CONSCIOUS DOUBLE VISION WHILE READING AND
READING AGE

	Percentage Numbers Showing Abnormalities		χ^2	P< (One tailed)
	Good Readers	Poor Readers		
Double vision while reading	13.3	44	4.02	0.5

To enable the pupil to understand more readily what is meant by double vision, a simple device designed for this project by Grant (1979) was used. This consisted of two identical transparent slides on which were printed a simple sentence. When the slides were in line the words were clear and identification was easy. As the slides were pulled apart, one part of a letter or word commenced to overlap a dissimilar part of the word beneath producing visual confusion. As the slides were pulled apart further the words and letters then appeared double.

Dynamic viewing

The dynamic binocular behaviour of the eyes was studied while actually moving and changing fixation during reading. Reading involves co-ordinating the eyes while scanning words and lines of words, and is thus a dynamic and not a static visual task. The head and eyes were observed without head restraint by a video camera, and the output recorded on a video recorder.

To assess the ability to co-ordinate the eyes on a non-reading task the pupils were asked first to change fixation between two crosses on a horizontal line 10 cm apart. In particular the purpose was to check scanning ability across the width of an average page of print.

Binocular eye behaviour while reading aloud an easy test passage of print was then observed. The test passage chosen was Test 2, form C, of Neale's Analysis of reading ability (Neale, 1958). The reading level of the passage was intentionally much below that of all the subjects (7 : 5 years).

The passage was placed at an angle of 10° to a horizontal table surface, and a video camera was used to monitor eye behaviour. The rating and interpretation of dynamic ocular behaviour while reading requires considerable experience and careful judgment. To ensure as reliable a rating as possible the subjects in both groups were placed in random order, and the video recordings, without sound, were viewed and assessed independently by observers experienced in this type of work. The variables considered were those found to be significant from previous research projects and from several years' experience investigating individual dynamic visual performance. In view of the difficulty involved in judging simultaneous but different movements of the eyes the levels of inter-rater reliability shown in Table 6 are reasonable. Their composite judged scores on the variables selected are also shown.

TABLE 6
RELATIONSHIP BETWEEN DYNAMIC VIEWING BEHAVIOUR AND READING AGE

Variable	Inter-rater Reliability Between Judges	Average Combined Ratings		P< (Mann Whitney U Test) Comparison One tailed
		Poor Readers N = 25	Good Readers N = 15	
Instability in co-ordinating the eyes on looking from left to right	0.748	3.56	2.67	0.01
Indecision as to controlling eye while reading	0.655	3.56	2.60	0.05
Unequal eye movements while reading	0.710	5.04	4.13	0.01
Both eyes tending to diverge while reading	0.795	5.24	4.20	0.001
One eye tending to diverge and then over-converge while reading	0.765	3.32	2.33	0.05
Compensatory head movements while reading	0.853	4.96	3.73	0.001
Facial muscular stress while reading	0.751	3.24	2.66	NS (0.06)

So that the assessments here should be on observations of visual factors only, no sound was recorded on the video tape, and therefore no oral clue given as to reading ability. The sound of the reading was recorded and analysed independently.

The results shown in Table 6 indicate that all but one of the variables identified are significantly associated with reading difficulty.

DISCUSSION

Static viewing

In the experiment on the relevance of letter chart visual acuity to reading difficulty no significant result was found. The purpose of this type of vision test is to establish normal visual acuity, and whether treatment is necessary (e.g., spectacles) to correct reduced vision. The visual and ocular factors involved in static viewing at distance are quite different from the dynamic vision involved while reading. The lack of significance of reduced visual acuity in relation to reading difficulty agrees with the results found by other researchers (e.g., Farris, 1934).

In the case of the tests on binocularity using the vision screener, the lack of significance of imbalances of the extra-ocular muscles assessed by dissociation tests both for distance and for near fixation in relation to reading difficulty confirms the work of other researchers referred to earlier. In this case binocularity only under static vision is being examined with visual displays quite different from those involved in reading, where both eyes are looking at the same display and scanning dynamically.

Stereoscopic vision is not generally examined as a routine, but involves a different aspect of binocular function than does a simple dissociation test of extra-ocular muscle imbalance; it also requires a higher order of binocularity. Reduced stereoscopic vision indicates reduced binocular function, and reduced ability to combine the images formed on each retina into an unconfused single visual percept. Though reading does not involve detection of difference in spatial position, it does appear that a reduced binocular function, as indicated by reduced stereoscopic visual acuity, using the Titmus Vision Screener, is significant in relation to reading difficulty.

In the case of the experiments on image suppression and reduced range of fusion for convergence and divergence, using polarised vectograms, the data did not indicate any significant relationship with reading difficulty. Apart from any difference in the type of visual display, it is also necessary to remember that the type and speed of movement of the vectograph display presented before each eye is likely to be different from that involved during reading. Theoretically the targets being viewed should be being fixated centrally on the fovea of each eye until the eyes can no longer linearly track the targets (when the images may gradually 'slip' across the retina) and fusion is no longer possible.

The range of fusion is usually assessed by gradually moving in and out two similar displays until they blur and/or go double. There is usually time during the test for a neural adjustment to be made, with the images slipping to some degree across the retina, before the subject notices a visual change, and can no longer perceive an unconfused image. Probably because of this situation, the range of fusion alone has been found to be statistically of minor significance in relation to reading difficulty. At this stage, however, the level of binocular investigation becomes somewhat more complex, and it is usually necessary to take into account a number of other related factors such as type of display, visual span on fixation, and dynamic visual behaviour, before a conclusion can be made in an individual case.

The experiment on the presence of conscious double vision while reading indicates one of the more obvious subjective manifestations of binocular image embarrassment that could be noticed, and could contribute to anomalies of visual perception and reading difficulty. Under dynamic viewing conditions double vision can occur on excessive divergence and over-convergence relative to the directions of gaze of the eyes required to maintain single vision.

Dynamic viewing

The studies of various aspects of dynamic binocular viewing, both while changing fixation and actually reading, indicate that as a more complex view is taken of visual behaviour, and the investigation is undertaken under the conditions operating for the particular visual task concerned, so very much more relevant data can be obtained.

Dynamic binocular co-ordination

Instability in co-ordinating the eyes on looking consecutively at two points spaced 10 cm apart has a high significance in relation to reading difficulty, and demonstrates likely difficulty in binocular control during scanning.

Indecision of controlling eye

Indecision as to the controlling eye while reading again indicates an unstable system of binocular co-ordination, and is likely to produce anomalies of visual perception and visual mis-cueing. The controlling eye in the context of this paper is taken as the eye which is most likely to maintain fixation centrally on to the fovea and macula area of most distinct vision on the retina.

Unequal eye movements

If the eyes move unequally while reading, and beyond the limits of individual corresponding retinal areas and of fusion, binocular image embarrassment and visual mis-cueing are likely to result. Therefore this variable is significant in relation to reading difficulty.

Divergence

The tendency for both eyes to diverge while reading is very significant in relation to reading difficulty, and is an indication that the child may be unable to converge adequately under dynamic viewing conditions, although convergence as assessed by the usual type of advancing static test may appear 'normal'.

Divergence and over-convergence

A tendency to diverge, and then over-converge, while reading is again a manifestation of difficulty in maintaining convergence under dynamic conditions, but here the degree of convergence effort is at times excessive, leading to temporary over-convergence. This cannot then be maintained and the eye diverges again.

Compensatory head movement

Variations of head movement are an attempt to compensate posturally for difficulties in scanning adequately ocularly, and appear to be highly significant in relation to difficulty in binocular co-ordination and to reading difficulty.

Muscular stress

The above studies on difficulties in co-ordinating the eyes during dynamic viewing, and indecision of ocular control while reading, indicate that the pupil must be under general muscular and nervous stress in attempting to compensate inadequately for these weaknesses. It is therefore reasonable that various signs of ocular and facial stress, such as frowning, screwing up the eyes, and excessive blinking do occur during reading, though the effect is not statistically significant for the short period of time of the test.

CONCLUSION

This paper has been concerned with the significance of anomalies of binocular vision under dynamic viewing conditions in relation to reading difficulty compared with a simpler investigation of vision and binocularity under static viewing conditions. Though a full discussion of the significance of the various anomalous retinal and cerebral stimulation situations that can arise from these anomalies of dynamic viewing is outside the scope of this paper it may be helpful to indicate what should be considered.

Though we are born with the facility of binocular vision, appropriate usage from birth is necessary to develop and adapt this function. If stable co-ordinated gaze of the eyes is not achieved, then the images of any object viewed will not regularly form on the same corresponding retinal areas, and permit full binocular function to develop. In other studies Bedwell (1973) has found that achievement of stable co-ordination often becomes more difficult on tasks, such as reading, which involve dynamic viewing on convergence. When images fall on retinal areas that have not developed binocular correspondence confused visual input can result. Manifestations of binocular image confusion are likely to be produced that may be associated with such situations as errors in reading through visual mis-cueing and difficulty in spelling because of poor visual sequential memory.

In the context of reading difficulty it is likely that the children who have never developed adequate ocular control will have difficulty in learning to read and to spell,

especially if taught by near visual methods. Others, who have learnt the elements of reading when young, especially if aided by phonics and some distance viewing of words, may later experience difficulty and possibly regression, if binocular control is inadequate to prevent confused visual input as the speed of work and visual fatigue increases.

ACKNOWLEDGMENTS—The authors would like to acknowledge the help given by the Rank Benevolent Fund towards the cost of the development of the visual technology required for this work and the interest and technical help given in the analysis of preparation of the data involved in this project by Mrs J. Barnes.

REFERENCES

- BEDWELL, C. H. (1973). The significance of visual ocular and postural anomalies in reading and writing. U.K.R.A. Conference. In *The Road to Effective Reading*, pp. 73-81. London: Ward Lock.
- BEDWELL, C. H. (1978). Binocular vision and near visual performance and fatigue. *Proc. 2nd World Congress Ergophthalmology*, pp. 383-397. Stockholm: Societas Ergophthalmologica Internationalis.
- BIRNBAUM, P. R. (1968). Binocular co-ordination as a factor in reading achievement. *Am. J. Optom. Ass.*, 39, 48-56.
- CRITCHLEY, M. (1964). *Developmental Dyslexia*. London: Heinemann.
- FARRIS, C. P. (1934). Visual defects as factors influencing achievement in reading. *California Sec. Educ.*, 10 (Oct.), 50-15.
- FLAX, N. (1968). Visual function in dyslexia. *Am. J. Optom.*, 45, 574-587.
- FLETCHER, M. C. (1969). The visual mechanism in learning disorders. *Am. J. Orthopt.*, 19, 92-96.
- GOLDBERG, M., and GUTHRIE, J. T. (1972). Problems in relating visual function to reading disorder. *Am. J. Optom.*, 47, 366.
- GOLDBERG, H., SHIFFMAN, N., and RUTTER, M. (1970). Evaluation of visual perceptual factors to reading disability. *J. pediat. Ophthalm.*, 9, 18-25.
- GOODMAN, K. S. (1975). Influences of the visual peripheral field in reading. *Res. Teach. English*, 9, 210-222.
- GRANT, R. (1979). *Classroom Tests for Binocular Instability*. Bognor: V.T.P. Aids.
- HINSHELWOOD, J. (1917). *Congenital Word Blindness*. London: H. K. Lewis.
- HIRSCH, S. M., and ANDERSON, R. P. (1976). Effects of motor training in reading achievement. In *Research Application in Reading*, pp. 163-181. Springfield, Ill.: Thomas.
- KEENEY, A. H. and KEENEY, V. T. (1968). *Dyslexia*. St. Louis: Mosby.
- NAIDOO, S. (1972). *Specific Dyslexia*. London: Pitman.
- NEALE, M. D. (1958). *Neale Analysis of Reading Ability*. London: Macmillan.
- RAYNER, K., and McCONKIE, G. W. (1977). *A Psychology of Reading*. London: Wiley.
- SCHONELL, F. J., and SCHONELL, D. E. (1950). *Diagnostic and Attainment Testing*. Edinburgh: Oliver and Boyd.
- SPACHE, G. (1976). *Diagnosis and Correcting Reading Difficulties*. New York: Allyn and Baum.
- TAYLOR, E. A. (1966). *The Fundamental Reading Skill*. Springfield, Ill.: Thomas.
- WHEELER, J. J., WATKINS, E. J., and McLAUGHLIN, S. P. (1977). Reading retardation and cross-laterality in relation to short-term information processing. *Br. J. educ. Psychol.*, 47, 126-131.

(Manuscript received 13th September, 1978)

RESEARCH NOTES

Br. J. educ. Psychol., 50, 71-73, 1980

EXTRAVERSION, NEUROTICISM, INTELLIGENCE AND ACADEMIC ACHIEVEMENT IN NORTHERN NIGERIA

By M. MAQSUD

(*Faculty of Education, Bayero University, Kano*)

SUMMARY. This study in the Northern Nigerian educational system tested the hypothesis of British and American studies that academic achievement at primary school is positively associated with extraversion. One hundred and sixty primary school boys completed the JEPI and Raven's Progressive Matrices. The relationships between the subjects' E, N, intelligence and achievement scores were examined. Contrary to the findings of British and American studies, the results of this enquiry yielded a significant negative association between extraversion and academic attainment.

INTRODUCTION

During the last two decades, evidence continued to accumulate that in addition to intellectual ability, there are some personality variables (e.g. extraversion, neuroticism) which significantly correlate with scholastic achievement. As to extraversion, a number of investigators found a low but significant positive association with scholastic achievement of primary school children (Jones, 1960; Morrison *et al.*, 1965; Rushton, 1966; Savage, 1966; Ridding, 1967; Entwistle and Cunningham, 1968). With regard to neuroticism, the findings of several studies indicated that among primary school children, high N scores are linked to poor academic performance (Astington, 1960; Lunzer, 1960; Hallworth, 1961; Callard and Goodfellow, 1962; Butcher *et al.*, 1963; Rushton, 1966; Savage, 1966; Entwistle and Cunningham, 1968).

Eysenck and Cookson (1969) pointed out that most of the above investigators employed the statistical technique of product-moment to analyse their data and paid very little attention to examining relationships between interactions of personality variables and academic attainment. These authors also suggested that "boys and girls may show quite different regressions of achievement or ability on personality variables" (p. 111) and thus they emphasised the need for treating the sexes separately. In his extensive review of the literature, Entwistle (1972) suggested several variables (sex, ability, type of subjects studied, classroom organisation and size, teaching methods, and geographical area) which may have interactions with scholastic achievement and personality variables. Since some interacting variables, such as teaching methods and schooling conditions, differ from one educational set-up to another and most of the above studies were conducted in Western settings where schooling conditions are more or less similar, one may expect different patterns of relationships between personality and academic success in some non-Western developing countries.

THE STUDY

It was decided to examine the relationships between personality, intelligence, and academic achievement at primary school age in a Northern Nigerian context.

Subjects: The 160 Class 7 boys in this sample were randomly drawn from two Government-owned primary schools in Kano City, a commercial and industrial town in Northern Nigeria. Their age range varied between 12 years 4 months to 13 years 7 months with an average of 13 years 2 months. The subjects belonged to lower-middle class homes and all of them took the Common Entrance Examinations for admission into secondary schools.

Tests: The Junior Eysenck Personality Inventory (JEPI) was used to take measures of extraversion and neuroticism, and Raven's Standard Progressive Matrices (RSPM) was administered to assess the subjects' non-verbal reasoning ability; their academic attainment in English and arithmetic was assessed by their marks, secured out of a maximum of 50 in each subject, for the Common Entrance Examinations.

Design and Statistical Analysis: There were three independent variables (extraversion, neuroticism, and non-verbal reasoning ability) and a dependent variable (academic attainment); all these variables entered into a $2 \times 2 \times 2$ factorial design. Analyses of variance were carried out for English and arithmetic scores separately.

RESULTS AND DISCUSSION

On the basis of E, N, and RSPM scores, the subjects were divided into two groups, high (Hi) and low (Lo) split at the median, for each of the three independent variables. Mean English and arithmetic scores with SDs for the six groups are given in Table 1.

TABLE 1
PERSONALITY, INTELLIGENCE AND ACADEMIC ACHIEVEMENT:
MEANS AND STANDARD DEVIATIONS

Variable	Group	N	English		Arithmetic	
			Mean	SD	Mean	SD
Extraversion	HiE (Mean = 12.33)	80	26.75	12.41	22.21	11.89
	LoE (Mean = 8.24)	80	31.59	11.70	27.30	10.74
Neuroticism	HiN (Mean = 15.69)	80	30.34	12.89	25.70	11.85
	LoN (Mean = 10.65)	80	28.00	11.56	23.81	11.28
Intelligence	HiIQ (Mean = 93.90)	80	33.40	10.40	28.40	9.89
	LoIQ (Mean = 75.10)	80	29.94	12.59	21.11	12.04

The data were analysed by means of a three-way analysis of variance. The analysis yielded significant main effects for extraversion, $F(1,152)=7.28$, $P<0.01$ for English and $F(1,152)=8.75$, $P<0.01$ for arithmetic, as well as for non-verbal reasoning ability, $F(1,152)=22.28$, $P<0.001$ for English and $F(1,152)=17.96$, $P<0.001$ for arithmetic. It was seen that RSPM scores were the most significant source of variance. Analyses of variance neither indicated significant main effect for neuroticism nor any significant interaction between the three independent variables.

Contrary to the findings of studies carried out in Western countries, the present results showed a significant negative association between E and achievement scores (i.e. introverts do significantly better than extroverts) and no significant link between N scores and attainment of Nigerian primary school children. These data are, however, in line with Entwistle's (1972) suggestion that there are several interacting variables, such as 'style of teaching', which may determine relationships between personality variables and academic attainment. It is interesting to note that the results are consistent for both English and arithmetic.

A significant negative association between extraversion and academic achievement of Nigerian primary school children may be attributed to teaching methods used in Nigerian schools. Pedagogical approaches employed in Nigerian primary schools lay heavy stress on serious academic concentration even in early primary classes, whereas primary school children in Western countries tend to employ more 'playful' learning activities. Though this explanation requires further empirical testing, it can tentatively be hypothesised that the nature of the relationship between personality and academic attainment mainly depends upon the character of curricular activities and teaching methods.

A significant positive association between intellectual ability and academic attainment of the subjects of this study substantiates the findings of earlier and recent studies and provides cross-cultural evidence.

REFERENCES

- ASTINGTON, E. (1960). Personality assessment and academic performance in a boys' secondary grammar school. *Br. J. educ. Psychol.*, 30, 225-236.
 BUTCHER, H. J., AINSWORTH, M. E., and NESBITT, J. E. (1963). Personality factors and school achievement—a comparison of British and American children. *Br. J. educ. Psychol.*, 33, 276-286.
 CALLARD, M. P., and GOODFELLOW, C. L. (1962). Three experiments using the Junior Maudsley Personality Inventory. Neuroticism and Extraversion in school boys as measured by JEPI. *Br. J. educ. Psychol.*, 32, 241-251.
 ENTWISTLE, N. J. (1972). Personality and academic attainment. *Br. J. educ. Psychol.*, 42, 137-151.

- ENTWISTLE, N. J., and CUNNINGHAM, S. (1968). Neuroticism and school attainment—a linear relationship? *Br. J. educ. Psychol.*, 38, 123-132.
- EYSENCK, H. J., and COOKSON, D. (1969). Personality in primary school children: 1—Ability and achievement. *Br. J. educ. Psychol.*, 39, 109-122.
- HALLWORTH, H. J. (1961). Anxiety in secondary school children. *Br. J. educ. Psychol.*, 31, 281-291.
- JENSEN, A. R. (1973). Personality and scholastic achievement in three ethnic groups. *Br. J. educ. Psychol.*, 43, 115-125.
- JONES, H. GWYNNE (1960). Relationship between personality and scholastic attainment. *Bull. Brit. Psychol. Soc.*, 40, 42.
- LUNZER, E. A. (1960). Aggressive and withdrawing children in the normal school. *Br. J. educ. Psychol.*, 30, 119-123.
- MORRISON, A., MACINTYRE, B., and SUTHERLAND, J. (1965). Teachers' personality ratings of pupils in Scottish primary schools. *Br. J. educ. Psychol.*, 35, 306-319.
- RIDDING, L. W. (1967). An investigation of the personality measures associated with over and under achievement in English and Arithmetic. *Br. J. educ. Psychol.*, 37, 397-398.
- RUSHTON, J. (1966). The relationship between personality characteristic and scholastic success in 11-year-old children. *Br. J. educ. Psychol.*, 36, 178-184.
- SAVAGE, R. D. (1966). Personality factors and academic attainment in junior school children. *Br. J. educ. Psychol.*, 36, 91-94.

(Manuscript received 15th August, 1979)

The British Journal of Medical Psychology

Volume 53, Part 1, March 1980

Selected contents

- C. Bagley. The factorial reliability of the Middlesex Hospital Questionnaire in normal subjects
- J. R. Bird. The borderline patient and military life
- A. K. Broome & S. M. Weaver. The generalization of new behaviours and maintenance of programme contingencies in a chronic locked ward
- M. Cramer. Psychopathology and shamanism in rural Mexico
- J. M. Damas Mora, F. A. Jenner & S. E. Eacott. On heautoscopy or the phenomenon of the double
- J. R. Korner. Phenomenological psychology and schizophrenia as explored by Erwin Strauss
- E. A. Marziali & J. M. Sullivan. Methodological issues in the content analysis of brief psychotherapy
- D. G. Millar. A repertory grid study of obsessionality
- D. Rampling. Written communications in psychotherapy
- N. D. Macaskill. The narcissistic core as a focus in the group therapy of the borderline patient
- R. W. Medlicott. The case of Joseph

Volume 53 (1980) £28.00 (US\$62.00). Single parts cost £9.00 (US\$20.00)

Volume 52 (1979) £21.00 (US\$45.00)

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN, UK

Br. J. educ. Psychol., 50, 74-76, 1980

CHILDREN'S EXPECTATIONS AND REALISATIONS OF SCIENCE

BY T. J. HARVEY AND PAULA EDWARDS

(*School of Education, University of Bath*)

SUMMARY. The evidence suggests that children in their first year of comprehensive school education in physics and biology did not display any increase in their interest in these subjects during the academic year. This is seen as particularly problematic for girls since they have, in any case, a low expectation in physics at the beginning of the year which was not affected by a year's experience in science.

INTRODUCTION

Why do so few women become scientists? Are they scientifically less able or are the pressures of society upon them so great that they are conditioned away from science at a very early age? One of the factors affecting future careers could be how interesting girls thought that science was going to be and whether it actually lived up to their expectations. The aim of this work was to see when young children's interests or expectations of science become fixed, and to see if their first experience at the secondary school altered their views.

Children in primary schools are taught very little, if any, science and it was thus hoped to develop an 'Interest in Science' or 'Expectation of Science' test for use with children in the age-range 9 to 12 years of age.

The vast majority of attitude-like tests demand that a person, having read a comment or statement, has to record his views on a five-point scale ranging from strongly agree to strongly disagree. The test thus relies upon the ability of the person to be able to read the initial statement and to comprehend the verbal choices offered. This approach may be suitable for skilled readers, but it seemed inappropriate for the age-range under test.

METHOD AND RESULTS

The research was thus conducted in four parts:

- (1) the production and testing of a suitable visual array to enable young children make choices ranging from 'strongly like' to 'strongly dislike';
- (2) the production of the interest test;
- (3) the application of the test to the relevant age-group;
- (4) the application of the test to the first year of a secondary school at the beginning and the end of the year.

Part 1. The testing of a visual array for recording preferences

Previous work by Harlen (1971) and Cooper (1976) had indicated that a series of circles could be used to represent a child's choice ranging from strongly like to strongly dislike. Various combinations of circles were tested upon groups of children in three local primary schools, the procedure adopted being a test, followed by an identical re-test four weeks later, and then correlating the two sets of responses. The final choice of the visual array was five circles increasing in size by equal increments of radius from left to right, with 'dislike' printed under the smallest, and 'like' under the largest. The children were asked to place a tick in the circle that best represented their degree of liking of the statement being tested.

Part 2. The production of the interest test

Forty-two science items containing 21 physical and 21 biological topics were selected. The children were asked to indicate on the five-point scale, by ticking the appropriate sized circle, how much they would like or dislike finding out about the items in question. The 42 items were given to 675 children, of which 201 were first year secondary, 247 were top primary and 227 were third year primary school children. The test was given during September and October and hence the secondary school children could still be considered to have had no formal science. Each item was scored on the scale 1-5, 5 being for 'like'.

If a single item is to be reliable the performance of the children upon it should be comparable with their performance upon the other similar items. The analysis was thus to correlate each item with the total score of the remaining 'like' items for each child. The physical science items were correlated together, as were the biological ones. The 12 items in each group having the largest item-to-total correlation were selected, giving a final test of 24 items.

Part 3. Sex and age difference in interest in science

The results of Table 1 show the mean scores and standard deviations per boy and girl for the final 24 items according to the three age groups.

TABLE 1
SEX DIFFERENCES IN INTEREST FOR SCIENCE BY AGE

Age Group	Sample Size		Physical Science				t-test	Biological Science					t-test
			Boys		Girls			Boys		Girls			
	Boys	Girls	Mean	SD	Mean	SD		Mean	SD	Mean	SD		
	9-10	120	107	41.7	11.8	32.4		12.0	39.5	9.2	41.9	10.9	
10-11	106	141	42.3	12.0	34.9	11.3	4.95**	39.7	10.5	42.6	9.6	2.24*	
11-12	107	94	43.8	10.4	30.8	10.5	8.84**	38.7	10.8	41.5	8.8	1.99*	

* $P < 0.05$

** $P < 0.01$

A t-test was performed upon the results to see whether the differences in the mean scores for the boys and girls were significant. The t-values obtained are also given in Table 1.

These results indicate the following:

- (i) Boys are significantly more interested in physical science than are girls.
- (ii) Girls are more interested in biological science than are boys, but the difference is not quite so marked as for physical science.

Since the two components of the test are not related to each other it is not possible to compare the relative interests in physical and biological science. In an attempt, however, to see how the interests changed with the age of the child an analysis of variance was performed upon the results. The resulting F-values are given in Table 2.

TABLE 2
ANALYSIS OF VARIANCE OF INTEREST CHANGES WITH AGE

	Physical Science		Biological Science		F-ratio for Significance	
	Boys	Girls	Boys	Girls	at 1%	at 5%
Between years	1.01	3.95	0.28	0.37	4.60	2.99

These results indicate that there is no significant difference between the interests of the children in any of the three age groups within the divisions shown. This supports the concept that once the child's interest has become fixed it will not alter providing there is no exposure to any formal science teaching. It is not clear at what age these interests are fixed other than before the age of nine.

Part 4. The effect of being taught science on a child's interest

The whole of a first year entry into a mixed comprehensive school were given the interest test at the beginning and end of their first year. The results are given in Table 3.

As expected, the results at the beginning of the year were almost identical with those obtained in Part 3 when the test was being developed and show the same sex/subject differences. A t-test was then carried out, the results being given in Table 3.

In all cases no significant difference between the interests of the children at the beginning and end of the year were found. Tragically it would appear that school has done nothing for them in terms of stimulating their interest in science.

TABLE 3
THE EFFECT OF BEING TAUGHT SCIENCE IN FIRST YEAR COMPREHENSIVE SCHOOL
BY SEX AND SCIENCE SUBJECT.

	Physical Science					Biological Science				
	Beginning of Year		End of Year		t-test	Beginning of Year		End of Year		t-test
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
Boys (N = 104)	43.5	9.70	44.6	9.3	1.88	36.9	10.8	36.6	10.8	0.20
Girls (N = 90)	31.7	9.43	29.9	9.3	0.58	38.6	9.6	40.3	9.0	1.61

CONCLUSION

During the first year at the secondary school the children were introduced to their first formal science. Their expectations seemed to have been realised since in no cases did the interest of groups significantly alter their results on the tests after a period of one school year. From a girl's point of view this is disastrous when considering physical science, for their initial expectation is very low and has not improved during the year.

Meyer and Penfold (1961) suggest that there is evidence to indicate that good attainment in science follows an interest in science. Kelly (1976) reports that boys perform significantly better in science than girls, and hence if one combines these two phenomena we see that girls relative to boys are very much disadvantaged and thus less likely to become physical scientists.

If secondary schools are unable to alter significantly children's interests in science and if once established these interests tend to remain fixed, the only solution appears to be to try and influence the children when they are actually becoming established. No research has yet been conducted in this area and so we remain ignorant as to when we can best help girls towards a favourable interest in physical science. Certainly, if we do not try to improve the situation for girls, they will remain disadvantaged and retain their present role as passive consumers of technology in an ever increasing technological society.

REFERENCES

- COOPER, C. (1976). The developments of electrical concepts in primary school children. M.Ed. Thesis, Bath University.
 HARLEN, W. (1971). *Report for Teachers on the Evaluation of the First Set of Units 5/13*. London: Schools Council.
 KELLY, A. (1976). Women in science: a bibliographic review. *Durham Research Review*, 36, 1092-1108.
 MEYER, G. R., and PENFOLD, D. (1961). Factors associated with interest in science. *Br. J. ed. Psychol.*, 31, 33-37.

(Manuscript received 15th November, 1978)

CONFERENCE ABSTRACTS

Br. J. educ. Psychol., 50, 77-93, 1980

Papers presented to the Education Section of the British Psychological Society at Oxford, 1979.

BAGLEY, C. R., D.Phil., Department of Sociology, University of Surrey, and
VERMA, G. K., Ph.D., A.B.P.S., Postgraduate School of Research in Education, University of Bradford.

THE STRUCTURES AND VALIDITY OF SELF-CONCEPT MEASURES IN MULTI-ETHNIC POPULATIONS

Building on our recent publications in this field (Bagley and Evan-Wong, 1975; Bagley and Mallick, 1978; Bagley, Mallick and Verma, 1979; Bagley, Verma, Mallick and Young, 1979) we report, first of all, comparisons of the factor structure of Coopersmith's self-esteem scale in comparable samples of West Indian, Asian and English pupils ($N=400$) drawn from a larger sample of adolescents attending multi-ethnic schools in England. Factor structure and item analyses show, in particular, different conceptions of self in West Indian males. In addition, between-sex comparisons indicate a different ranking of self-esteem in the West Indian group, in comparison with the English group. These results will be compared with factor analyses of Coopersmith self-esteem data, gathered in the course of a study of secondary education in India (Verma, Bagley and Mallick, 1979).

In the second part of this paper we will report studies of the comparative (face) validity of different methods of self-esteem in multi-ethnic middle schools, comparing short versions of the Coopersmith and Piers Harris self-concept measures with an adjective check list measure of self-characteristics, and a range of cognitive variables.

REFERENCES

- BAGLEY, C., and EVAN-WONG, L. (1975). Neuroticism and extroversion in responses to Coopersmith's self-esteem inventory. *Psychol. Rep.*, 36, 253-254.
BAGLEY, C., and MALLICK, K. (1978). Development of a short form of the Piers-Harris self-concept scale. *Educ. Rev.*, 30, 265-268.
BAGLEY, C., MALLICK, K., and VERMA, G. (1979). Pupil self-esteem: a study of black and white teenagers in British schools. In VERMA, G., and BAGLEY, C. (Eds.), *Race, Education and Identity*. London: Macmillan.
BAGLEY, C., VERMA, G., MALLICK, K., and YOUNG, L. (1979). *Personality, Self-Esteem and Prejudice*. Farnborough: Saxon House.
VERMA, G., BAGLEY, C., and MALLICK, K. (1979). *Illusion and Reality in Indian Secondary Education*, Farnborough: Saxon House.

BAMBER, J. H., and

BILL, J. M., Department of Education, The Queen's University of Belfast.

IN TWO MINDS—ARTS AND SCIENCE SPECIALISATION AT SIXTH-FORM LEVEL

In 1959 C. P. Snow's *The Two Cultures* focused attention on what he regarded as the increasing split between the literary and scientific cultures into two polar groups 'who had almost ceased to communicate at all, who in intellectual, moral and psychological climate had so little in common . . . one might have crossed an ocean'. Snow's pinning of the blame on the extreme specialisation of the British educational system has been followed by investigations at the sixth-form level of psychological and cultural differences between the arts and science specialists. Contrasts in intellectual and personality characteristics between the two specialisms have been shown by Hudson (1968) and by Child (1978). Though the prevailing notion of the science specialist as 'culturally ignorant' is not confirmed by Richmond (1963), the view of the scientist (*vis-à-vis* his arts peer) as 'illiterate' and lacking 'an educated vocabulary' is lent support by Kirkman (1964) and in AEB examiners' reports in chemistry and physics (1976). The investigations described in this paper examine further the contention that already by sixth form level arts and science specialists are polarised into two widely differing kinds of minds.

Two separate investigations of male and female sixth-form arts and science specialists (total N=522) were conducted in Northern Ireland schools (mixed and single-sex). The first of these studies examined performance on a variety of measures of literacy together with measures of culture, personality and intelligence. The second focused on skills in critical thinking alongside measures of culture, value-systems and intelligence.

Summarising the results of these two separate studies it was found that (1) scientists were no less literate than their arts counterparts on tests of mechanical accuracy and were actually superior on a number of measures of 'facility with words'; (2) compared with the arts specialists scientists showed up as 'more rounded' on measures of culture; (3) scientists outperformed arts specialists on measures of intelligence (including verbal) and on tests of critical thinking; (4) on personality dimensions, scientists emerged as more stable and more tough-minded than arts specialists; (5) scientists and arts specialists differed in value-systems with the former showing up as more 'theoretical' and more 'economic' and the latter as higher on aesthetic, social and religious values; (6) a number of pupil sex and school type effects also emerged.

The evidence of these two studies therefore (in line with Richmond, 1963) leads us to question the contemporary view of the scientist as 'culturally narrow' and relatively illiterate, while pointing to a number of potentially important psychological differences between arts and science specialists, including the hitherto neglected area of values. These differences, it is suggested, may have implications for the educational guidance of students, for the problems of early specialisation (is it the arts person who suffers more?) and, in so far as school sex type influences patterns of specialisation, for the 'co-ed vs single sex' debate.

REFERENCES

- Associated Examining Board (1976). Reports of Chemistry and Physics Examiners: Section 6, Science.
 CHILD, D. (1978). *Inaugural Lecture*, University of Newcastle.
 HUDSON, L. (1968). *Frames of Mind*. London: Methuen.
 KIRKMAN, A. J. (1964). Command of vocabulary among university entrants. *Educ. Res.*, 9, 151-159.
 RICHMOND, W. K. (1963). *Culture and General Education*. London: Methuen.
 SNOW, C. P. (1959). *The Two Cultures and the Scientific Revolution*. Cambridge: Cambridge University Press.
 Physics Examiners (1976). Reports of Examiners.

BAYNE, R., Ph.D., Recruitment Research Unit, Civil Service.

PEAK EXPERIENCE AND SELF DISCLOSURE

Responses to a peak experience questionnaire devised by Paffard (1970) were scored in two ways: (1) direct and similar to Paffard's own method and (2) an indirect content analysis. To compare the two methods of scoring, each was related to two possible measures of self-actualisation of 'openness to one's own experiencing': the Dogmatism scale and the Jourard Self-Disclosure Questionnaire (JSDQ). The same pattern of results was found in two studies, providing evidence *against* the direct method of scoring and the JSDQ, and *for* the indirect content analysis and the Dogmatism scale. The implications of the results for the measures, for education, and for Maslow's concept of peak experience are discussed.

REFERENCE

- PAFFARD, M. K. (1970). Creative activities and 'peak' experiences. *Br. J. educ. Psychol.*, 40, 283-290.

BREESE, J. H., M.A., University of London Goldsmiths' College.

TEACHING ABOUT INDIVIDUAL DIFFERENCES TO PROSPECTIVE AND PRACTISING TEACHERS

In their answers to the amount of importance to be attached to the aim of 'enabling student teachers to understand pupils' intellectual development' 65 out of 68 responding lecturers from university departments and colleges of education attributed considerable

importance to this aim (Breese, 1974). The writer went on to suggest that where the teaching of educational psychology is given a high priority, room might be found on a PGCE one-year course for a five-session course on individual differences. Freeman (1974) also listed individual differences as one of the topics covered in the applied psychology at Chorley College of Education.

In reality, the writer has only been able to devote one session rather than five to individual differences in the five-session course he has run in the psychology of learning, covering, with the help of handouts, points related to language, intellectual and personality development and differences. While the material relating to language and personality does not seem to have been particularly emotive, that relating to intellectual differences as measured by tests such as Raven's Matrices on occasions has roused strong feelings among a few students, particularly if the work of Jensen (1971) is mentioned.

The present paper will describe the material the writer uses and will aim to open up discussion on the material covered by other lecturers, the reactions of their students, and on how the topic of individual differences in intellectual ability can best be taught to prospective and practising teachers.

REFERENCES

- BREESE, J. (1974). The teaching of educational psychology to student teachers. *Psychol. Teaching*, 2, 22-43.
 FREEMAN, J. (1974). The case for applied psychology in education. *Psychol. Teaching*, 2, 43-50.
 JENSEN, A. R. (1971). Do schools cheat minority children? *Educ. Res.*, 14, 3-28.

BROWN, A. E., M.A., Ed.D., Co-ordinator of Student Services, Board of Education for the City of Toronto, Canada.

THE TORONTO EARLY IDENTIFICATION AND DEVELOPMENTAL PROGRAMME

In 1967 a mixed research and service programme was initiated by the Psychological Service of the Board of Education for the City of Toronto in an endeavour to develop a procedure for identifying children at risk educationally early in their school careers. The research phase of the project was terminated in 1974 by which time data had been collected, much of it longitudinal, on over 25,000 children in approximately one hundred schools.

Three aspects of the programme will be outlined:

- (1) A brief overview of the research resources with respect to the various approaches, procedures and instruments which were investigated.
- (2) Consideration of the evolution of the programme from being primarily a means of identifying children at risk to a psychological service delivery system primarily concerned with generating developmental data for use in school programming for all the children in a class.
- (3) A consideration of the political and professional issues involved in implementing a systems approach of this kind and implications for future development of the programme.

REFERENCES

- BROWN, A. E. (1976-77). Intelligence tests and the politics of school psychology. *Interchange*, 7 (3).
 BROWN, A. E., and PERKINS, M. J. (1977). The Toronto Early Identification and Developmental Programme: evolution and expansion. *The Ontario Psychologist*, Vol. 9, No. 4.
 LANDRUS, G. D., BROWN, A. E. and LONG, E. R. (1974). *The Toronto Early Identification and Developmental Programme*. Department of Psychological Services, Toronto Board of Education.

CHAPMAN, B. L. M., Ph.D., University of Bristol.

SOME IMPLICATIONS OF VARIABLE INTERACTION FOR MODELS OF TEACHING PROCESS

It is beginning to emerge from a small number of studies in school effectiveness, including one by the author, that a number of relationships between measures of school process which

are considered important are markedly curvilinear. The evidence of such relationships is not all recent but work in this area of investigation has hitherto been marked by a reluctance to depart from a strictly linear model and so many salient non-linearities have been missed.

Rosenshine (1971) proposed the existence of such relationships in a review of work relating teaching behaviour to student achievement and Preece (1976) cites both Thompson and Bowers (1968) and Nuttall (1974) as having found 'U' shaped relationships. Levy and Tucker (1972) report on a three-factor interaction relating streaming, sex and attainment while there is considerable *prima facie* evidence of interaction between the incidence of educational handicap and attainment from numerous studies of low achievers referred to centres specialising in educational handicap. This author (1979) reports on a re-analysis of a large data base which is yielding numerous strong interaction effects between criterion and explanatory variables.

Strong curvilinear effects in relation to teaching variables are amenable to a variety of explanations centred on the notion of saturation of teaching or matching of pace of presentation between teacher and pupil. The latter notion appeals to the concept of tracking and may be so represented in a simulation exercise.

Less easy to explain in these terms are the findings of Levy and Tucker relating to streaming and the finding of the writer that strong interactions are also found between attainment and a number of measures of pupil attitude including views on disciplinary matters, academic aspiration, expectation and administrative matters. A model which has considerable face validity and which goes some way towards explaining these effects regards the student as a component in a feedback loop in which rate of learning rather than quantity learned is the dominating factor. In principle the parameters of such a model are quantifiable and the transfer function will yield predictive data by means of which its validity may be tested.

It is too early yet to be confident that such an approach will provide an analytical tool for regular use but at least the predictions of such a model at a conceptual level are proving to be extremely useful in terms of the present data base.

REFERENCES

- CHAPMAN, B. L. M. (1979). Schools do make a difference. *Br. ed. Res. J.*, in press.
 LEVY, P. and TUCKER, J. (1972). Differential effects of streaming on primary school attainment. *Br. J. educ. Psychol.*, 42, 75-79.
 NUTTALL, G. A. (1974). Is classroom interaction research worth the effort involved? *N.Z. J. educ. Studies*, 9, 1-17.
 PREECE, P. F. W. (1976). Achievement and the interaction of teaching and learner variables. *Research Intelligence*, Vol. 2, No. 1.
 ROSENSHINE, N. (1971). *Teaching Behaviours and Student Achievement*. Slough: NFER.
 THOMPSON, G. R., and BOWERS, N. W. (1978). Fourth grade achievement as related to creativity, intelligence and teaching style. *American Educational Research Association*.

DAVEY, A. G., Ph.D., Reader in Applied Social Studies, University of Newcastle upon Tyne.

THE DEVELOPMENT OF RACIAL AWARENESS IN YOUNG CHILDREN

This four-year project, financed by the Social Science Research Council, began in 1975.

The study is designed to investigate the factors associated with the development of racial awareness in primary school children between the ages of seven and eleven. Samples of West Indian, Asian and White children have been drawn from areas of London and Yorkshire which differ in their socio-cultural traditions and their concentrations of different ethnic groups. In addition a 'no contact' sample has been selected from rural schools in North Yorkshire.

A variety of measures were constructed to investigate the children's perception of ethnic differences, the saliency of cues used as grouping criteria and their emotional response to people perceived as belonging to groups other than their own. A sociometric study has also been made of the children's friendship patterns and all the parents have been visited by experienced 'own race' interviewers. A selection from the principal findings of this investigation and their implications for educational practice will be given.

DRIVER, G., B.A., M.A., Ph.D., Centre for Applied Social Studies, University of Leeds.

ETHNICITY AND CULTURAL COMPETENCE: ASPECTS OF INTERACTION IN MULTI-RACIAL CLASSROOMS

Studies of the school performance of ethnic minority pupils have hinged around measurable constructs and test instruments devised in isolation from the subjects of research. The paper reviews such work associated with self-esteem, identity, self-control and intelligence. It goes on to consider such constructs and their status in the perspective of ethnomethodological accounts of classroom interaction.

The paper takes into account the writer's own ethnographic and statistical research into the school performance of West Indian boys and girls. The findings indicate that such constructs as self-esteem and identity are to be viewed as situationally variable. The critical contexts in which roles and relationships are defined are seen to have cultural and political dimensions which influence attitudes and shape outcomes in terms of achievement.

REFERENCES

- ADAMS, R. N. (1975). *Energy and Structure*. Austin: University of Texas Press.
 COOPERSMITH, S. (1968). *The Antecedents of Self Esteem*. San Francisco: Freeman.
 DRIVER, G. (1977). Cultural competence, social power and school achievement: West Indian pupils in the West Midlands. *New Community*, V (4), 353-359.
 DRIVER, G. (1979). *Good Black Marks*. London: Commission for Racial Equality.
 GOFFMAN, E. (1959). *The Presentation of Self in Everyday Life*. New York: Doubleday.
 GOODENOUGH, W. H. (1971). *Culture, Language and Society*. Reading, Mass.: Addison-Wesley.

DRIVER, ROSALIND, Ph.D., and

JOHNSON, SANDRA, M.Sc., Centre for Studies in Science Education, University of Leeds.

MONITORING SCIENCE PERFORMANCE: SOME DESIGN CONSIDERATIONS

The Assessment of Performance Unit, announced in 1974, is set up within the DES. Its terms of reference are:

"To promote the development of methods of assessing and monitoring the achievement of children at school, and to seek to identify the incidence of under-achievement."

The model adopted for monitoring is one based on lines of development each of which, it is argued, have relevance across subjects in the curriculum: pupils' scientific development¹ is one such 'line'.

As part of its brief, the science monitoring team² was given a list of criteria oriented to the assessment of scientific processes.

The framework and rationale³ for the types of tests being developed together with sample questions, will be presented.

There are a number of novel features associated with the monitoring programme. One of these is the 'process orientation' of the tests, and a second is the pupil/question sampling strategy to be applied. These have implications for necessary development work in the areas of question writing, test construction and data analysis. These and related issues will be outlined and discussed.

¹ This model was originated by Brian Kay, HMI in KAY, B. W. (1975), Monitoring pupils' performance. *Trends*, 2, 11-18.

² The science team, appointed in 1977, is based at Chelsea College, University of London, and the Centre for Studies in Science Education, University of Leeds.

³ See "Assessment of Performance Unit, Steering Group on Science: Progress Report 1977-78", DES, December, 1978.

GHUMAN, P. A. S., B.A., M.Ed., Ph.D., A.B.Ps.S., and
MOLYNEUX, A., B.A., M.Ed., The University College of Wales, Aberystwyth.

AN INQUIRY INTO THE USE OF SOME PSYCHOLOGICAL CONCEPTS AND
THE MANNER IN WHICH THEIR UNDERLYING CONNOTATIVE MEANINGS
ARE INTERPRETED BY THREE GROUPS OF EDUCATORS

In educational psychology stipulative definitions are often used in an attempt to be objective and value neutral, but these definitions may easily be misinterpreted by the classroom teachers and lecturers to be prescriptive and/or programmatic.

The present paper sets out to investigate empirically the manner in which certain psychological terms such as extravert, neurotic, needs, growth are given connotative meanings by two groups of teachers and a group of college lecturers: primary school, secondary school, college of education lecturers. Conceptual analysis of the psychological terms is not presented as it will constitute a paper in itself. (See Scheffler, 1960; Louch, 1966; Hirst and Peters, 1970; Peters, 1970, 1973.)

The following hypotheses were set up:

- (1) The lecturers would give different connotative meanings to some or all the psychological terms as compared with school teachers.
- (2) All the three groups would perceive certain concepts, e.g., extravert, stable and motivation, more favourably than their counterparts' concepts.

In all 147 subjects were tested: 47 primary school teachers; 50 secondary school teachers and 50 lecturers in education. No claim is made that the sample was a random one; on the contrary several institutions (27 in all) were contacted, and questionnaire was only given to those individuals who showed willingness to co-operate. Data were collected by using the semantic differential scale with its three major dimensions: evaluation, potency, and activity.

The data were analysed by using the t-test and one way ANOVA. The major finding is that the three groups do perceive the concepts differently, i.e., accord different connotative meanings, especially on the evaluation dimension. There is a small but positive indication that the lecturers rate some concepts neutrally as compared with the two groups of teachers, which substantiates, partly, our first hypothesis. There is also support for the second hypothesis—though not as clear-cut as we would have liked—that no matter what the academic psychologists say about the value neutrality of the concepts (hypothetical or otherwise) (Eysenck, 1971), the psychological terms such as introvert, neurotic and stable, as far as evidence from this sample is concerned, acquire emotive values and lose their objectivity. If this were the case universally it could have profound effects on labelling and stereotyping of children, which in turn could affect their academic performance and behaviour in the classroom.

REFERENCES

- EYSENCK, H. J. (1971). *Race, Intelligence and Education*. London: Temple-Smith.
HIRST, P. H., and PETERS, R. S. (1970). *The Logic of Education*. London: Routledge and Kegan Paul.
LOUCH, A. R. (1966). *Explanation and Human Action*. California: University of California.
PETERS, R. S. (Ed.) (1973). *The Philosophy of Education*. Oxford: Oxford University Press.
PETERS, R. S. (1970). *The Concept of Motivation*. London: Routledge and Kegan Paul.
SCHEFFLER, I. (1960). *The Language of Education*. Springfield, Illinois: Thomas.

GURNEY, P., M.Ed., School of Education, University of Exeter.

THE USE OF BEHAVIOUR MODIFICATION TO ENHANCE SELF-ESTEEM
IN MALADJUSTED BOYS

This research investigation attempted to examine the association between the frequency of positive self-referent verbal behaviour in maladjusted boys and their level of self-esteem, as reflected both in overt behaviour and in self-report measures. Twenty-seven ten- to twelve-year-old maladjusted boys were pretested on self-esteem and 14 low scorers were randomly allocated to experimental and control groups.

In the main experimental periods all boys were interviewed daily by the experimenter to discuss their school work and diaries, and to play a scoring game. The behaviour modification programme, designed to increase positive self-referent verbal statements, followed an ABCDA design. Baseline was followed by a three-part treatment 'package'.

Target behaviour increased substantially in the experimental period and was maintained at a high level. It was hypothesised that at post-test the experimental group would receive significantly higher ratings of overt behaviours related to self-esteem than the control group, and secure significantly higher scores on self-report tests of self-esteem, locus of control and anxiety. A significant difference was found in the staff ratings of overt behaviours related to self-esteem. A follow-up study carried out after a further interval of five weeks did not detect any additional significant differences.

REFERENCES

- AVILA, D., and PURKEY, W. (1972). Self-theory and behaviourism: a rapprochement. *Psychology in the Schools*, 9, 124-126.
 HAUSERMAN, N., *et al.* (1976). A behavioural approach to changing self-concept in elementary school children. *Psychological Record*, 26, 111-116.
 PARKER, H. C. (1974). Contingency management and concomitant changes in elementary school students' self-concepts. *Psychology in the Schools*, 11, 70-79.

HAHN, H., M.A., Paris Business School, Oxford.

THE DEVELOPMENT OF AN EDUCATIONAL PSYCHOLOGIST'S ROLE IN AN INNOVATIVE EUROPEAN BUSINESS SCHOOL

Stemming from a French initiative, 45-70 young people aged 20-25 years participate in a French-British-German educational programme which involves living, working and studying in three European countries and functioning effectively in three languages. The development of the project from its inception in 1973 is outlined. This will include an explanation of the pedagogical principles of the school based on the concepts of self-responsibility, action-reflection, and learning from experience, and these will be illustrated by reference to the work of the staff and students. The paper then focuses on the role of the 'moniteur' in the United Kingdom—a role which was initiated in France, and developed in the United Kingdom from the base of an educational and career guidance model as practised at the Tavistock Institute of Human Relations. The way in which a psychologist-client model has changed into a staff-member-student working relationship is described and illustrated by reference to both theory and practice.

In addition to reporting on the way in which a professional psychological approach has been integrated into a business educational programme, the paper reflects on issues relating to the inter-dependence of staff roles, aspects of the developmental process in young adults, and processes of cross-cultural learning.

HORNSBY, B., M.Sc., M.Ed., L.C.S.T., University of North Wales, Bangor.

EVALUATING THE EFFECT OF TEACHING DYSLLEXIC CHILDREN

Investigations over the past 12 years have shown that the prognosis for all backward readers is poor (Morris, 1966; Clark, 1970; Rutter *et al.*, 1970; Start and Wells, 1972) and that for specific reading retardates (dyslexics) is even worse, in spite of their overall higher intelligence level (Yule, 1973; Warnock, 1978).

The question of approach to the teaching of reading has also received considerable attention (Southgate, 1970; Smith, 1971).

The aim of the experiment was to demonstrate that retarded/discrepant children can be helped, rather than to compare the differences in results between the centres chosen to be studied. Three such centres where the concept of dyslexia is taken seriously were selected, firstly because their backgrounds were so different—one being a hospital clinic supported entirely by the National Health, the second a peripatetic service attached to a university and the third a private fee-paying establishment; and secondly because their methods of treatment are so similar, are published and can be studied. The centres also run training courses from which their personnel are selected.

Since control groups are neither ethical nor really possible in a study of this kind, the subjects are acting as their own controls, by comparing the *ratio of gain to time taught before* treatment began to the *ratio of gain to time taught after* treatment began. The null hypothesis would be that there is no difference, but in fact a considerable difference was found and discussion attempted as to why this should be.

REFERENCES

- CLARK, M. (1970). *Reading Difficulties in Schools*. London: Penguin Papers in Education.
 MORRIS, J. (1966). *Standards and Progress in Reading*. Slough: N.F.E.R.
 RUTTER, M., WHITMORE, K., and TIZARD, J. (1970). *Education, Health and Behaviour*. London: Longmans.
 SMITH, F. (1971). *Understanding Reading*. New York: Holt, Rinehart and Winston.
 SOUTHGATE, V. (1970). *Reading—Which Approach?* London: English University Press.
 START, K. B., and WELLS, B. K. (1972). *The Trend of Reading Standards*. Slough: N.F.E.R.
 WARNOCK REPORT (1978). *Special Educational Needs*. London: H.M.S.O.
 YULE, W. (1973). Differential prognosis of reading backwardness and specific reading retardation. *Br. J. educ. Psychol.*, 43, 244-248.

KYRIACOU, C., University of Cambridge, Department of Education.

ROTTER'S I-E SCALE: A PRINCIPAL COMPONENTS ANALYSIS

The paper presents an analysis of data obtained from 130 secondary school teachers who completed a Likert scale version of Rotter's Internal-External locus of control scale. The results indicated that teachers who were younger, less experienced, and university graduates appeared to have a greater belief in external control compared with their respective subgroups of colleagues. A principal components analysis revealed four factors which appeared to underlie the scale. However, following a discussion of comparisons between the present factor structure and factor structures reported in other studies, it is argued that the possible multidimensionality of the scale does not undermine its use.

LINDSAY, G. A., B.Sc., Cert.Ed., M.Ed.(Ed.Psych.), Ph.D., A.B.Ps.S., Sheffield Education Department, Psychological Service; and Sheffield University, Institute of Education.

EARLY IDENTIFICATION OF LEARNING DIFFICULTIES: A CRITIQUE

Screening children for learning difficulties has become a popular process. There has been a change in emphasis away from screening for attainment (normally reading) to screening children on abilities hypothesised as pre-requisite for later academic success. Thus, children have been screened at earlier ages, and the 'at risk' concept has come to the fore.

It will be argued that these attempts at early screening have not, in general, been investigated as rigorously as is desirable; that when tests and instruments are investigated most are found to be flawed; and that even the best available instruments are usually presented in the literature in an optimistic fashion.

Reference will be made to work in Sheffield over the past four years, including the validation of the Infant Rating Scale, as a means of early identification of learning difficulties, and the joint working of psychologists and advisers on an in-service education programme considering the monitoring of children's development.

Finally a way forward based upon in-service education of teachers rather than screening *per se* will be advocated.

REFERENCES

- CORNWALL, K. (1979). The development of a screening technique: some reflections on four years experience in one local authority. D.E.C.P. Occasional Papers, 3, 3-10.
 LINDSAY, G. A. (1979). The early identification of learning difficulties and the monitoring of children's progress. Unpublished Ph.D. thesis, University of Birmingham.
 SATZ, P., and FLETCHER, J. M. (1979). Early screening tests: uses and abuses. *J. Learn. Dis.*, 12, 56-60.

MERRETT, F., M.A., B.Sc. (Econ.), Dip.Ed., Faculty of Education, Wolverhampton Polytechnic.

TEACHING TEACHERS TO USE BEHAVIOUR MODIFICATION IN THE CLASSROOM

The psychology taught in colleges of education and university departments tends not to be sympathetic to the behavioural viewpoint. Descriptive child development and/or odd lectures on motivation, memory or personality are considered acceptable but there is usually nothing to form a basis upon which the technology of behaviour modification could rest. Likewise, teaching method lectures are largely concerned with scheme and lesson preparation and specific skill or subject teaching and pay little attention to general issues of class management.

First, a small survey was made among primary class teachers in a West Midlands inner ring area which showed clearly that many teachers were well aware of the difficulties of class management and pointed to some specific problems. Next, a small project was set up to see if a behavioural intervention could be operated by an intelligent and well-motivated teacher with some guidance but with no previous knowledge. This was successful (see Merrett and Wheldall, 1978). Two attempts were then made to see whether students in training could benefit from a knowledge of behaviour modification when on practical teaching assignments and whether their attitudes towards the technology would change as a result. The practical problems involved in the application of behaviour modification by students in this situation proved to be considerable but attitudes were shown to change. This suggested that courses for students in training could help to bring about a more sympathetic attitude towards behaviour modification but that in-service courses for practising teachers were likely to be the best way of testing the usefulness of the techniques and of winning more popular acceptance of them.

An in-service course held at a teachers' centre has met with considerable interest in schools and great enthusiasm from those taking part. The teaching methods which will prove to be best in enabling the teachers to apply behavioural methods with conviction and to achieve success in the classroom are still being investigated. One great strength of the in-service course is the opportunity it provides to try out some ideas in the classroom and then have the opportunity to discuss the outcome with other interested and informed people at intervals. A positive feature of this first in-service course has been the arrangement of periodic workshop sessions for this purpose now that the course is complete.

REFERENCE

MERRETT, F., and WHELDALL, K. (1978). Playing the game; a behavioural approach to classroom management. *Educ. Rev.*, 30, 41-50.

Moss, G., M.Ed. (Ed. Psych.), M.Ed., LL.B., D.C.P., D.Comm., A.B.Ps.S., Avon Educational Psychology Service.

EDUCATIONAL PSYCHOLOGISTS AND IN-SERVICE TRAINING FOR TEACHERS: DEVELOPING TEACHER WORKSHOPS IN BEHAVIOURAL PSYCHOLOGY

The conventional relationship of educational psychologist to teacher is through the referral of a 'problem pupil'. This approach inevitably focuses more upon the defects of the child as a learner than upon the deficiencies of the adult as a teacher—the pathological model of intervention. The competence of this model to meet most teachers' needs is being increasingly questioned by today's educational psychologists (Gillham, 1978) who are seeking more effective involvement in the development and evaluation of the in-service training of teachers (Moss, 1979). The sort of psychology that some educational psychologists may dispense in their INSET contribution is often still of the descriptive kind, e.g., general cognitive or emotional development in children. However, some psychologists have been developing course-based approaches that intend some practical and immediate relevance to teachers' daily needs through the techniques of behavioural psychology. Behavioural psychology applied to classroom practice may focus upon 'problem behaviour management'

within the ordinary school (e.g., Merrett and Wheldall, 1978), or upon task analysis within the special school (e.g., Cocks and Gardner, 1979).

This paper describes the development and content of teacher workshops utilising behavioural psychology as a model for analysing the conditions under which classroom learning may take place more efficiently. Because of its stress upon the precise analysis of teacher-pupil behaviour, the method is more frequently known as precision teaching, and derives from the pioneering work of Ogden Lindsley. Although precision teaching may be only one of several formulations for changing behaviour, it has in common with all behaviour modification procedures that firstly the problem is defined in the operational terms of the behaviour itself, secondly the 'entering behaviour' is measured and recorded, thirdly some procedure for modifying that behaviour is introduced while continuing to record the behaviour, and fourthly the analysis and intervention procedure is revised as necessary.

REFERENCES

- COCKS, R., and GARDNER, J. (1979). The role of the educational psychologist in the education of severely subnormal children. *J. Assoc. Educ. Psychologists*, 4, 10.
 GILLHAM, B. (Ed.) (1978). *Reconstructing Educational Psychology*. London: Croom Helm.
 MERRETT, F., and WHELDALL, K. (1978). Playing the game: a behavioural approach to classroom management. *Educ. Rev.*, 30, 41-50.
 MOSS, G. (1979). Bringing educational psychologists within the network. *Insight: Journal of the National Conference of T.C. Leaders*, 3, 1.

PICKTHORNE, B., A.C.P., Ph.D., and

WHELDALL, K., B.A., Ph.D., A.B.Ps.S., Department of Educational Psychology, University of Birmingham.

MOTIVATING THE UNMOTIVATED: A BEHAVIOURAL APPROACH TO TEACHING HIGHER EDUCATION STUDENTS

Problems of poor motivation and consequent unsatisfactory attainment in subsidiary and ancillary courses in higher education are ubiquitous. The fact that many such courses are compulsory only exacerbates the problem. This paper reports an appraisal of such a situation and describes a behavioural technique which proved to be successful in resolving it efficiently. Technical and organisational difficulties which often arise at this level of education were solved by the use of a modified form of Keller's Personalised System of Instruction (Keller, 1968; Keller and Sherman, 1974). The students were given a special 'learning package' consisting of a series of notes, data sheets, problems and laboratory experiments. A number of precise learning objectives for each topic were specified and students were given 'feedback tests' on the topics at two-weekly intervals. Students who successfully reached criterion or 'target score' (which was successively raised by small increments) were allowed an afternoon break and to finish the tutorial session early. Those who failed to meet the criterion had to undertake remedial work at these times, thus giving them a dual incentive to work harder in future. Students reported that the clear goals and regular attainment tests, coupled with the personalised learning format (which enabled them to work at their own pace), provided a clear picture of their progress and reinforced their desire to learn. This was reflected in significantly improved scores on the criterion tests over the period and also in other assessments of their progress on the course.

REFERENCES

- KELLER, F. S. (1968). Goodbye teacher. *J. Appl. Behav. Analysis*, 1, 79-80.
 KELLER, F. S., and SHERMAN, J. G. (1974). *The Keller Plan Handbook*. Menlo Park, Calif.: W. A. Benjamin Inc.

PUMFREY, P.D., A.B.Ps.S., Division of Educational Guidance, Department of Education, University of Manchester.

SOME SELF-PERCEIVED PROBLEMS EXPERIENCED BY ADOLESCENTS: IDENTIFICATION AND INTERVENTION

Increasing attention is being paid in our schools to identifying various difficulties experienced by adolescent pupils. The development of school-based services enabling pupils

to deal constructively with such problems is also taking place. The idea that each school should have an integrated guidance programme within which competent personnel can provide both individual and group counselling is gaining greater acceptance (Scottish Education Department, 1968; Antonouris, 1974; Chamberlain and Delaney, 1977).

In this paper the development, uses and limitations of a two-section self-report inventory devised for use with third-year pupils attending secondary schools will be outlined (Lewis and Pumfrey, 1978). The first section of the inventory is designed to elicit the self-perceived problems of adolescents in six relatively discrete areas of experience and to differentiate between pupils in terms of the number of self-perceived problems acknowledged. The second section of the inventory is intended to encourage a more detailed exploration of the individual student's self-perceived current difficulties.

Statistical and clinical evidence concerning the validities and reliabilities of the instrument based on various researches will be presented. Recent findings concerning the relationship between locus of control belief and self-perceived problems will also be reported (Aspin, 1979). Some implications of these results will be considered within the context of a school guidance programme.

REFERENCES

- ANTONOURIS, G. (1974). Subsequent careers of teachers trained as counsellors. *Br. J. Guidance and Counselling*, 2, 160-170.
- ASPIN, G. (1979). Locus of control belief and the self-perceived problems of adolescents. Unpublished Diploma in Guidance and Counselling dissertation, University of Manchester, Department of Education.
- CHAMBERLAIN, J., and DELANEY, O. (1977). Guidance and counselling in Irish schools. *Br. J. Guidance and Counselling*, 5, 49-54.
- LEWIS, D. G., and PUMFREY, P. D. (1978). *The Lewis Counselling Inventory*. Slough: N.F.E.R.
- Scottish Education Department (1968). *Guidance in Scottish Secondary Schools*. Edinburgh: H.M.S.O.

RAVEN, J. Scottish Council for Research in Education.

BARRIERS TO EDUCATION AND EDUCATIONAL INNOVATION FROM OUTSIDE THE SCHOOL SYSTEM

Data will be presented to support the following contentions:

(1) Parents want their children to develop markedly different, indeed incompatible, qualities. Nevertheless, they are opposed to different children being catered for in different ways. This is because they fear that, if different children were treated in different ways, the most articulate and the most powerful would get the best deal. It follows from this that one cannot help pupils to develop value-laden competencies because helping them to do so would involve helping them to develop these qualities in relation to their own valued goals. It also follows that one cannot recognise, foster, or reward individual characteristics. Finally, it follows that one cannot have parental involvement in schools because different parents would make different, and irreconcilable, demands on teachers.

(2) Most parents think that secondary education is primarily about getting jobs, not about developing the competence of their children. When asked what would happen if their children did not do well at school, most said they would fail to get good jobs, and only a minority said that their intellectual and other abilities would fail to develop.

(3) What actually happens in school classrooms is determined by what is assessed in the scramble for jobs, not by the wishes and priorities of pupils, parents, employers, and employees. If one wishes to introduce curriculum innovations it is therefore necessary to gain control over what is assessed in examinations. This is the primary task, not a secondary one.

(4) While most teachers would welcome the opportunity to give their pupils credit for qualities not assessed in examinations, most are extremely reluctant to see the credentialing system extended to encompass assessments of the goals which they themselves believe to be important. The reasons for this resistance are therefore crucial to the improvement of educational practice, and some of the reasons for this resistance will be explored.

RENNIE, E. F. N., B.Sc., Dip.Ed., M.Sc., A.B.Ps.S., Kirklees School Psychological Service.

THE WEST RIDING SCREENING SIX YEARS ON

In 1973 the former West Riding County Council carried out a comprehensive screening procedure of all the children then in top infants. Various aspects of this screening survey have been previously reported (e.g., Pickles, 1974; Rennie, 1975). However, the present study differs in that it is a follow-up after six years of the children identified as having difficulties as infants in 1973.

The survey was a two-stage process and resulted in the identified children being divided into four groups. In the present follow-up the subsequent reading progress of these groups is compared, both between themselves and with a normal population. The lists of identified children have also been compared with the lists of those on roll in special schools and those referred to the school psychological service from the relevant age group. It is concluded that whole age group educational screening on this model is probably not cost-effective. Another interesting finding is that of the children failing to read five words or more on the Burt Graded Word Reading Test at age seven the majority now score within normal limits on a reading comprehension test.

REFERENCES

- PICKLES, D. G. (1974). The Burt (Rearranged) Graded Word Reading Test: Percentile Scores for 27,000 West Riding Children of Final Infant Age. N.F.E.R. Test Information Sheet 45:02.
 PICKLES, D. G. (1974). Adjustment Problems in the Infant School: The Use of Stott's Six Adjustment Pointers for Screening Purposes. Unpublished Manuscript 1974.
 RENNIÉ, E. F. N. (1975). Reproduction of designs in early screening. *J. Assoc. educ. Psych.*, Vol. 3., 10, 30-33.

RYAN, J. P., University of Manchester.

THE FACTORIAL STRUCTURE OF THE RUTTER QUESTIONNAIRE FOR TEACHERS

Completed Rutter B(2) questionnaires for 1257 children were examined using a series of principal components factor analyses rotated to the Varimax and then to the Promax position ($K=4$). The discussions centre on the Promax factor patterns. A full sample analysis yielded eight factors of which four appeared to be important and descriptive of maladjustment. These were named as Antisocial, Neurotic, Overactive and Rejection, and were associated with the same second-order factor. Three lesser factors seemed to describe conditions associated with maladjustment (Absenteeism, Oral Immaturity, and possibly Neurological) and a final factor expressed mainly residual variance. Separate analyses for boys ($N=676$) and girls ($N=581$) gave similar but not identical patterns. For girls the factors were very much like those for the full sample, but for the boys the eighth factor emerged fairly strongly as Disturbed Delinquent. Comparison of the Antisocial and Neurotic factors for the sexes with the similarly named 'sub-scores' of the scale shows as many differences as agreements and the question arises as to whether the factorially-based or the 'discriminative ability' based set should be used.

Separate age analyses (but with sexes combined) were carried out on four age groups: $4\frac{1}{2}+$ ($N=149$), $7\frac{1}{2}+$ ($N=409$), $10\frac{1}{2}+$ ($N=442$), $13\frac{1}{2}+$ ($N=257$). Factors were restricted to eight at each age level although nine had occurred at both the lowest age groups on a preliminary analysis. Overall the factor pattern was roughly the same, but the factor content in terms of items loading upon them showed a definite developmental progression for both the major factors, Antisocial and Neurotic. In particular the $4\frac{1}{2}+$ group had the Antisocial factor split into two: F3 = Antisocial (fractious) and F7 = Antisocial (delinquent), which is what would be expected of the age group. An expected change also occurred with the $13\frac{1}{2}+$ group where the Rejection factor combined with the Neurotic to give a picture of a more adult kind of neurotic pattern. At the same time the Absenteeism factor became more closely associated with the other maladjusted factors.

The implications for changes in the scale, for changes in the scale norms, and the provision of separate norms at age x sex levels will be discussed.

SELBY, C. H., HMI, Assessment of Performance Unit, Department of Education and Science, London.

ASSESSMENT OF PERFORMANCE—WHY, WHAT AND HOW?

The assessment of Performance Unit (APU) was announced in a White Paper (Command 5720) in 1974 and set up in 1975 together with the Educational Disadvantage Unit. The terms of reference of the APU are: "To promote the development of methods of assessing and monitoring the achievement of children at school and to seek to identify the incidence of under-achievement." It has been recognised that the design and development of suitable test instruments requires careful study and close consultation with all concerned in their administration and interpretation and that these necessary procedures take time.

So far, monitoring programmes have been funded for mathematics, language and science and assessment is involving pupils aged 10+ and 15+ in all three areas together with the 13+ age group in the case of science. There are plans to monitor the first foreign modern language also at 13+ but no date has been given for the start of this particular programme. The monitoring teams for mathematics and languages are situated at the National Foundation for Educational Research and have benefited from the earlier studies of tests of attainment of mathematics in schools (TAMS) and the work of the Bullock Committee respectively. Science monitoring is being shared by teams at the Leeds University School of Education and the Centre for Science Education, Chelsea College, University of London. Assessment procedures in mathematics and science include practical exercises as well as pencil and paper tests. In language as well as reading and writing there will be a study of speaking and listening skills. In addition children's attitudes towards various aspects of language and mathematics are being reported on.

The APU has set up a statistics advisory group to give advice on sampling, monitoring and item banking in the various aspects of the curriculum under review. Only a small sample of pupils is tested in any monitoring survey and because of this and the confidential nature of the procedures used the reports will not identify the individual performances of pupils in schools or local education authority areas. Monitoring surveys in mathematics began in the spring of 1978 and those in language in the spring of 1979 and surveys in science are expected to start in 1980. It is expected that, in the first instance, monitoring will continue each year at each age range in each of these three areas.

The first report in each area of the curriculum will provide a broad picture of performance. Subsequent reports will help to give an increasingly detailed picture and also to trace changes in performance over time. There should be sufficient detail to identify aspects of performance where it would seem appropriate to mount more concentrated studies. The overall aim will be to contribute to the background against which decisions about the allocation of resources are taken and to provide the basis for well informed debate about educational performance.

SUMNER, R., M.Ed., Ph.D., A.B.Ps.S., N.F.E.R., Slough.

MONITORING PERFORMANCE VARIATIONS

The APU has carried out its first surveys in Mathematics at 11+ and 15+, and in Language (Reading and Writing) at 11+. Thus it has established data which will form the basis for future comparisons with data from subsequent surveys. Conventionally, the appraisal of 'change' requires the comparison of different sample means obtained at two or more points of time when any effects attributable to survey timing on the samples' ages are allowed for (e.g., DES, 1978).

Moving away from the use of a set-piece test to a more complete representation of the curriculum raises a number of strategic questions. The methods adopted by the USA National Assessment Programme (NAEP publications) will be described. These will be contrasted with the possibilities available to the APU in the context of its terms of reference and the types of assessment materials currently in use or being developed. A preliminary discussion may be found in the proceedings of a Council of Europe Contact Workshop (Sumner, 1977).

The paper will be speculative (a) because the APU is itself considering the issues against a background of a phased introduction of monitoring in three curriculum areas, viz. Mathematics, Language and Science, and (b) the author is not involved either at the NFER or the APU with the formulation of long-term policies. Any views expressed will be his own.



REFERENCES

- SUMMER, R. (Ed.) (1977). *Monitoring National Standards of Attainment in Schools*. Proceedings of a Council of Europe Contact Workshop, N.F.E.R., Slough.
- D.E.S. (1978). *Primary Education in England: A Survey by H.M. Inspector of Schools*. London: H.M.S.O.
- NAEP—See Publications list issued by National Assessment Publication, Education Commission of the States, 1860 Lincoln Street, Suite 700, Denver, Colorado 80203.

THOMAS, T. C. H., Dyfed.

EARLY IDENTIFICATION OF LEARNING DIFFICULTIES AND PRACTICAL STRATEGIES OF INTERVENTION

Children were screened on the Swansea Infant Evaluation Profile. The children who were identified as being educationally 'at risk' were the subject of an intensive intervention programme. The educational impact of this pattern of intervention was evaluated when the control group was tested at the end of their infant school careers. The results suggest that the experimental 'at risk' group did better than was predicted for them while the control 'at risk' group did less well than was predicted for them. Early analysis of the follow-up suggests that the experimental 'at risk' children have sustained their rate of progress after the cessation of the intervention programme whereas the control group have done less well. This is in spite of the evidence from the Swansea Profile which suggested that the control 'risk' group had greater potential for educational achievement than could be expected for the experimental 'risk' group.

Two hundred and nineteen children were involved in this study which was financed by the Welsh Education Office and was based at a local teachers' centre.

TOMLINSON, SALLY, M.Soc.Sci., Ph.D., Department of Educational Research, University of Lancaster.

WEST INDIAN CHILDREN AND ESN-M SCHOOLING—A CASE OF NON-ACHIEVEMENT

Although special (ESN-M) education is usually presented as catering for the 'needs' of certain children, to be categorised out of the normal education system, in a society which values qualifications, represents the ultimate in non-achievement. Since the development of schools and classes for 'educable defective' children in the 1890s there have been individual protests but no articulate group protest at classification. The protest of the West Indian community, that 'too many' of their children are referred and ascertained as ESN-M, represents the first articulate protest to arise on a group level. The 'ESN issue' has taken on a profound significance in that it now symbolises the general under-achievement of children of West Indian origin in the school system.

A recent small-scale study of the decisions made on 40 children passing through the referral and assessment process (18 of the children being of immigrant parentage) is used to suggest explanations for the over-referral of West Indian children. Heads of referring schools, educational psychologists, medical officers and heads of special schools were asked their views on the 'problems' of West Indian children—these were then compared with their general accounts of how they judged children to be ESN-M. Referred criteria and assessment judgements corresponded very closely to the 'problems' West Indian children were considered to have. In contrast, Asian children did not meet the referral criteria, and this may help to explain the small numbers of Asian children in ESN-M schooling.

Finally, the serious implications for the schooling and educational under-achievement of children of West Indian origin, of the suggested Warnock category "Child with Learning Difficulty" is discussed in this paper.

REFERENCES

- TOMLINSON, S. (1978). West Indian children and ESN schooling. *New Community*, VI, No. 3.
- TOMLINSON, S. (1979). Decision-making in Special Education, (ESN-M) with some reference to the Children of Immigrant Parentage. Unpublished Ph.D. thesis, University of Warwick.

TURNBULL, H., M.A., M.Ed., Ph.D., A.B.Ps.S. and,
DAVIDSON, G., B.Sc., M.Ed., A.B.Ps.S., Craigie College of Education, Ayr.

A REGIONAL SURVEY OF ATTAINMENT IN READING AND MATHEMATICS

The aim of this survey was to ascertain levels of performance in reading and mathematics in a representative sample of primary school children in the Dumfries and Galloway Region. This was to extend to consideration of individual as well as overall levels and to performance on individual items on the tests and the various sections of the reading tests.

The three age levels were 7 : 3 to 8 : 2, 9 : 6 to 10 : 5 and 10 : 6 to 11 : 6. As the Region requested that the sample be a large one, some 25 per cent of each age group were tested. The total numbers involved were approximately 2,400, with some 800 in each group. Thirty-one schools, ranging from urban to small rural, were involved. One of the most important aspects of the work, the analysis of which is still ongoing, is to meet the requirements of the individual schools as well as the Region. Consideration of the tests to be used called for consultation amongst all the people involved, including the educational psychologists. As a result, Stages 1, 2 and 3 of the Edinburgh Reading Tests were chosen along with Basic Mathematics Tests A and DE and the Staffordshire Test of Computation.

Costing was an important consideration at this stage and procedures had to tie in with the availability of funding, met partly by the Region and partly by the Scottish Education Department.

The arrangements for administering and scoring the tests were made at a meeting of all the head teachers and their psychologists. The latter, along with the college staff, instructed the teachers in the test procedures. Scoring of tests was carried out by staff of different schools.

Coding each individual test for computer analysis followed and this proved to be a lengthy procedure. Details of the data to be processed on the SPSS package and information required had been discussed with the staff of the Computer Education Department at Jordanhill College of Education.

Results, which are still being produced, have shown a satisfactory overall level and information from the grids of the Basic Mathematical Tests and the various sections of the Edinburgh Reading Tests have been informative. Intercorrelations amongst the sections of the Reading Tests and between them and the Basic Mathematics Tests have been relatively high.

REFERENCES

- Basic Mathematics Tests A (1975) and DE (1974). N.F.E.R.
Edinburgh Reading Tests, Stages 1, 2 and 3. London: Hodder and Stoughton.
HEBRON, M. E., and PATTINSON, W. (1973). Staffordshire Test of Computation. London: Harrap.
Stirlingshire County Council Education Committee (1974). *Report of the Investigation into Standards in Primary Schools*.

WATSON, G., B.A., Department of Psychology, University of Reading.

SEX ATTITUDES, KNOWLEDGE AND EDUCATION OF ESN(M) CHILDREN

ESN(M) children tend to come from lower socio-economic groups, and Wilson (1973) has described an element of political conservatism within such groups. Sexually, however, Eysenck (1976) has suggested that they are permissive. Nevertheless, Hall *et al.* (1973) found a tendency for the mildly retarded to be conservative in their sexual attitudes.

Two surveys are reported. In the first, 194 ESN(M) children were given a questionnaire to ascertain their level of sexual knowledge and attitudes, and results were compared with a control group of 61 children. The ESN(M) children were found to have a significantly lower level of knowledge than the control children, and their attitudes were significantly more traditional. Children who reported having learned about sex from a teacher scored higher than those who had learned from home or peers. In addition a test/retest of children taking a school sex education course showed gains in knowledge and a shift towards more liberal attitudes.

In the second survey 231 (62.0 per cent response) questionnaires were analysed from ESN(M) secondary schools in England and Wales and compared with 104 (51.7 per cent)

questionnaires from mainstream control schools. The ESN schools reported giving more time to sex education and regarded a wider range of topics as important. They also showed more staff involvement and a greater use of the vernacular during instruction. At the same time, in comparison with the control schools, they showed a preference for topics concerning relationships and family life rather than those with a factual content such as contraception and sexually transmitted disease.

REFERENCES

- EYSENCK, H. J. (1976). *Sex and Personality*. London: Open Books.
 HALL, J. E., MORRIS, H. L., and BARKER, H. R. (1973). Sexual knowledge and attitudes of mentally retarded adolescents. *Am. J. ment. Defic.*, 77, 706-709.
 WILSON, G. D. (1973). *The Psychology of Conservatism*. London: Academic Press.

WEDELL, K., Ph.D., University of Birmingham.

EARLY IDENTIFICATION AND COMPENSATORY INTERACTION

Current concern with the early identification of children with learning problems derives from two main sources: the findings of research on developmental learning difficulties in children, and the findings of research on the effects of experiential deprivation. Both these areas of research have demonstrated a relationship between later educational achievement and the early stages of children's development, but both have failed to demonstrate a one-to-one correspondence in the case of individual children. Attempts to identify children 'at risk' have involved two kinds of indicators—signs of later problems and early manifestations of later problems. Attempts to evaluate these indicators have tended not to substantiate their predictive validity, and this seems to be due to a simplistic view of the causation of educational success and failure. It is suggested that the scope for prediction is limited by the compensatory interaction between the resources and deficiencies within the child and within his environment, and this is illustrated by some experimental evidence.

It is suggested that the monitoring of children's progress is of more use than attempts to predict it. Projects which help teachers achieve clearer and more detailed performance objectives provide opportunities for such monitoring, and enable them also to identify those children who need more resources than the ordinary classroom can provide. Teachers also need help in maintaining an appropriate level of expectation for their children. While these measures may help children who have started at school, the needs of those below this age should be tackled with similar strategies.

WHELDALL, K., B.A., Ph.D., A.B.Ps.S., and
 CONGREVE, SUE, B.Phil.(Ed.), Department of Educational Psychology, University of Birmingham.

TEACHERS AND BEHAVIOUR MODIFICATION:
WHAT DO THEY THINK OF IT SO FAR?

Behavioural approaches to teaching have become a major growth industry in the United States but have been slow to catch on in the UK. Unlike their North American counterparts, British teachers are largely ignorant of behavioural techniques since only a minority will have encountered them during the course of their work or in their college training. A subjective estimate of British teachers' attitudes towards behaviour modification was of mild antipathy based on scanty knowledge. No empirical evidence regarding this was apparent in the literature leading to the present study which employed Musgrove's (1974) scale to measure attitudes towards behaviour modification.

One hundred and sixteen mature students following advanced courses in education anonymously completed Musgrove's scale assessing their attitudes towards behaviour modification and its use in educational settings. Their overall mean score suggested a general ambivalence of attitude, this sample, as a whole, being neither strongly for nor against behaviour modification. Neither sex, age, number of years' teaching experience nor age or type of children taught was significantly related to attitude score, but there was strong evidence to suggest a strong positive relationship between prior knowledge of behaviour modification by study or practical experience and attitude score. The finding that the more

teachers knew about behaviour modification the more positive were their attitudes was also supported by a retest study. Forty-four of the original sample were retested after completing a course in behaviour modification, resulting in significantly more positive attitude scores. Factor analyses provided support for Musgrove's original findings.

REFERENCE

MUSGROVE, W. (1974). A scale to measure attitudes toward behaviour modification. *Psychology in the Schools*, 2, 68-71.

WOLFENDALE, SHEILA, Croydon School Psychological Service.

IDENTIFICATION OF EDUCATIONAL PROBLEMS: THE CROYDON MODEL

Various influences can be traced that have had bearing on the acknowledged growth area of the 1970s, known generally as early identification screening for learning difficulties (or educational problems, or children educationally 'at risk'). A trend parallel to developments in the UK can be seen in America, Canada, Australia and parts of Europe.

The theoretical, conceptual and methodological issues pertaining to hypothesis testing in a service setting such as education are currently being explored in various parts of the UK. The main concerns focus around forms of validity and reliability of various measures, whether or not prediction can be an educationally valid concept, forms and feasibility of follow-up intervention, and the role of the teacher.

The model of developmental ('profile') screening adopted in the London Borough of Croydon, fully, from 1972, has epitomised the commitment of the LEA to a philosophical stand, acknowledging:

- (1) the central role of the class teacher and the need to sharpen and develop professional expertise in child development and the recognition of special educational needs;
- (2) the need for schools, with the support of educational agencies, to evolve their own coherent policies for the teaching of reading and for 'remedial' intervention;
- (3) the necessity of providing varying forms of in-service training, and constantly re-examining the rationale and *raison d'être* of the system.

As far as psychological perceptions and involvement have been concerned, the following aspects have been explored within the framework of an early identification procedure: reliability of the measure according to teachers; the pragmatic usefulness of the measure in giving teachers information as a basis for further action; forms of intervention, and ways of monitoring progress; feasibility of a pre-school link-up; pre-referral features of learning difficulty and learning stress.

REFERENCES

A video film for in-service teacher training on the rationale of early identification screening, and the origins of learning failure, followed by examples of school-based action has been made for general use in LEAs and College Departments of Education. *Early Screening Procedures*, D.E.A. Teacher Training Series, published by Drake Educational Associates Ltd., 212 Whitchurch Road, Cardiff, CF4 3XF. A course handbook accompanies the series.

The main text is *Identification of Learning Difficulties—a Model for Intervention* by Sheila Wolfendale and Trevor Bryans, 1979, published by the National Association for Remedial Education (N.A.R.E.) 2 Lichfield Road, Stafford.

PUBLICATIONS RECEIVED

Books for review have been selected from the following ('n.p.' means 'no price stated').

- ABBS, P. (1979). *Reclamations: Essays on Culture, Mass-Culture and the Curriculum*. London: Academic Press, pp. viii+134, c. £5.95.
- AINSCOW, M., and TWEDDLE, D. A. (1979). *Preventing Classroom Failure: an Objective Approach*. Chichester: John Wiley and Sons, pp. x+205, c. £7.50.
- BARKER, P. (1979). *Basic Child Psychiatry*. St. Albans: Granada, pp. xiv+274, c. £10.00.
- BERKE, J. H. (1979). *I Haven't Had to go Mad Here*. Harmondsworth: Penguin Books, pp. 218, p. £1.25.
- BORNSTEIN, M. H., and KESSEN, W. (Eds.) (1979). *Psychological Development from Infancy: Image to Intention*. London: John Wiley and Sons, pp. v+404, c. £10.00.
- BOWER, G. H. (Ed.) (1978). *The Psychology of Learning and Motivation*. London: Academic Press, pp. xiii+263, n.p.
- BROWN, B. B., and BEVERIDGE, M. (1979). *Language Disorders in Children: a Workshop for Teachers*. London: The College of Speech Therapists, pp. 68, n.p.
- BURNS, R. B. (1979). *The Self Concept Theory, Measurement, Development and Behaviour*. Harlow: Longman, pp. ix+341, p. £6.50.
- CADE, C. M., and COXHEAD, N. (1979). *The Awakened Mind*. London: Wildwood House, pp. v+275, p. £3.95.
- CASHDAN, A. (1979). *Language, Reading and Learning*. Oxford: Blackwell, pp. vii+171, c. £8.95, p. £2.95.
- CATANIA, C. A. (1979). *Learning*. New Jersey: Prentice-Hall, pp. v+406, c. £11.00.
- CRAIG, G. J. (1979). *Child Development*. New Jersey: Prentice-Hall, pp. xiii+524, n.p.
- DE LEMOS, M. M. (1979). *Aboriginal Students in Victoria*. Victoria: ACER, pp. xii+195, p. \$10.00.
- DOMAN, G. (1979). *Teach Your Baby Maths—Kit*. London: Jonathan Cape, p. £4.95.
- DOWNNEY, M., and KELLY, A. V. (1979). *Theory and Practice of Education—an Introduction* (2nd Ed.). London: Harper and Row, pp. xv+323, p. £3.50, c. £6.95.
- DYER, J. R. (1979). *Understanding and Evaluating Educational Research*. Reading, Mass.: Addison-Wesley, pp. xvi+430, c. \$14.95.
- FORSYTH, J. P., and DOCKRELL, W. B. (1979). *Curriculum and Assessment: the Response to Munn and Dunning—a pre-Publication Summary*. Edinburgh: SCRE, pp. 18, n.p.
- FOUCAULT, M. (1979). *Discipline and Punish*. Harmondsworth: Penguin Books, pp. 333, p. £2.95.
- FREEMAN, J. (1979). *Gifted Children*. Lancaster: Falcon House, pp. xii+294, c. £6.95.
- FURNEAUX, B., and ROBERTS, B. (Eds.) (1979). *Autistic Children—Teaching, Community and Research Approaches*. Henley on Thames: Routledge and Kegan Paul, pp. xiv+193, p. £2.75.
- GALLOWAY, D. M., and GOODWIN, C. (1979). *Educating Slow Learning and Maladjusted Children*. Harlow: Longman, pp. xii+163, p. £3.95.
- GATHORNE-HARDY, J. (1979). *The Public School Phenomenon*. Harmondsworth: Penguin Books, pp. 7+511, p. £2.95.
- GOLDSTEIN, H. (1979). *The Design and Analysis of Longitudinal Studies—Their Role in the Measurement of Change*. London: Academic Press, pp. v+199, c. £11.80.
- HAMLIN, D. W. (1979). *Experience and the Growth of Understanding*. Henley on Thames: Routledge and Kegan Paul, pp. xi+159, p. £2.75.
- HAMLIN, D. W. (1979). *The Psychology of Perception—A Philosophical Examination of Gestalt Theory and Derivative Theories of Perception*. Henley on Thames: Routledge and Kegan Paul, pp. vii+120, p. £2.25.
- HENSHAW, A. (1979). *After Sixteen (Full-time Further Education for Mentally Handicapped Students at Herefordshire College of Art and Design)*. London: NSMHC, pp. v+88, p. £2.50.
- JOHNSON, B. (1979). *Reading Appraisal Guide*. Victoria: ACER, pp. iv+79, p. \$4.50.
- KAUFMAN, A. S., and WEINER, I. B. (1979). *Intelligent Testing with the WISC-R*. Canada: John Wiley and Sons, pp. vii+268, c. £12.25.
- KAUFMAN, L. (1979). *Perception—The World Transformed*. Oxford: Oxford University Press, pp. viii+416, c. £10.75.
- KELLY, J. G. (Ed.) (1979). *Adolescent Boys in High School (A Psychological Study of Coping and Adaptation)*. New Jersey: John Wiley and Sons, pp. vii+267, c. £10.00.

- KENNY, D. A. (1979). *Correlation and Causality*. Canada: John Wiley and Sons, pp. vii+277, c. £12.20.
- KLIMA, G., and BELLUGI, U. (1979). *The Signs of Language*. London: Harvard, pp. xii+417, c. £16.25.
- KLINE, P. (1979). *Psychometrics and Psychology*. London: Academic Press, pp. v+380, c. £19.00/\$40.00.
- KOZLOFF, M. A. (1979). *A Program for Families of Children with Learning and Behaviour Problems*. London: John Wiley and Sons, pp. ix+450, c. £14.00.
- LAGAN, J. (1979). *The Four Dimensional Concepts of Psycho-Analysis*. Harmondsworth: Penguin, pp. xi+290, p. £2.95.
- LANGFORD, P. (1979). *Beyond Piaget—Recent Theories of Concept Development and their Significance for Teaching*. Australia: La Trobe University, pp. 110, p. \$6.00/£3.50.
- LIEBERT, B., and LIEBERT, M. (1979). *A Schoolwide Secondary Reading Program: Here's How*. Chichester: John Wiley and Sons, pp. x+402, c. £7.50.
- LUBIN, G. L., POULSEN, M. K., MAGARY, J. F., and SOTO-MCALISTER, M. (Eds.) (1979). *Piagetian Theory and the Helping Professions*. (7th Annual Conference, Vol. II.) California: University Affiliated Program, University of S. California, pp. 1+466, n.p.
- MACKAY, C., and COX, T. (1979). *Response to Stress—Occupational Aspects*. Surrey: IPC Science and Technology Press, pp. vii+224, p. £10.00/\$26.00.
- MARJORIBANKS, K. (1979). *Families and their Learning Environments—An Empirical Analysis*. Henley on Thames: Routledge and Kegan Paul, pp. xiii+255, c. £7.95.
- MCDONNELL, J. H. (1979). *Children's Riddling*. Bloomington: Indiana University Press, pp. xii+212, p. \$17.50.
- MEDCOF, J., and ROTH, J. (Eds.) (1979). *Approaches to Psychology*. Milton Keynes: Open University Press, pp. vi+354, p. £4.95.
- MEYER, M. E. (1979). *Foundations of Contemporary Psychology*. Oxford: Oxford University Press, pp. v+726, c. £12.00.
- MITTLER, P. (1979). *People Not Patients—Problems and Policies in Mental Handicap*. Hampshire: Methuen, pp. viii+238, p. £4.50, c. £7.95.
- MORGAN, G. (1979). *A Criterion-referenced Measurement Model with Corrections for Guessing and Carelessness*. Victoria: ACER, pp. v+76, p. \$5.00.
- MORRISON, C. M. (1979). *Open Plan Secondary Schools—An Annotated Bibliography*. Edinburgh: SCRE, pp. 26, n.p.
- MUSSEN, P. (1979). *The Psychological Development of the Child* (3rd Ed.). Hemel Hempstead: Prentice-Hall, pp. xiv+127, p. £3.85, c. £7.10.
- MUSSEN, P. H., CONGER, J. J., and KAGAN, J. (1979). *Child Development and Personality* (5th Ed.). London: Harper and Row, pp. v+579, p. £6.95, c. £9.70.
- MUSSEN, P. H., CONGER, J. J., KAGAN, K., and GEIOWITZ, J. (1979). *Psychological Development: A Life-Span Approach*. London: Harper and Row, pp. iv+502, c. £8.00.
- NAVARICK, D. J. (1979). *Principles of Learning: From Laboratory to Field*. Oxford: Addison-Wesley, pp. vi+436, n.p.
- OLWEUS, D. (1978). *Aggression in the Schools. Bullies and Whipping Boys*. Chichester: John Wiley and Sons, pp. xiii+218, c. £9.95.
- O'NEIL, H. F. (Jr.) (1979). *Procedures for Instructional Systems Development*. New York: Academic Press, pp. ix+321, c. \$19.50.
- O'NEIL, H. F. (Jr.) (1979). *Issues in Instructional Systems Developments*. New York: Academic Press, pp. v+211, c. \$17.50.
- PHYE, G. D., and RESCHLY, D. J. (Eds.) (1979). *School Psychology—Perspectives and Issues*. New York: Academic Press, pp. v+349, c. \$22.00.
- PYLE, D. W. (1979). *Intelligence—an Introduction*. Henley on Thames: Routledge and Kegan Paul, pp. vii+123, c. £5.95, p. £2.95.
- RAGGETT, M., TUTT, C., and RAGGETT, P. (1979). *Assessment and Testing of Reading. Problems and Practices*. London: Ward Lock Education, pp. 207, p. £4.95.
- ROBINSON, N. M., ROBINSON, H. B., DARLING, M. A., and HOCH, G. (1979). *A World of Children—Daycare and Preschool Institutions*. California: Brooks/Cole, pp. xii+250, p. \$8.95.
- ROGERS, T., and CAUSLEY, C. (1979). *Those First Affections—An Anthology of Poems Composed Between the Age of Two and Eight*. Henley on Thames: Routledge and Kegan Paul, pp. xiv+150, c. £5.95, p. £2.95.

- RYRIE, A. C., FURST, A., and LAUDER, M. (1979). *Choices and Chances*. London: Hodder and Stoughton for SCRE (pub. 71), pp. 140, n.p.
- SAUNDERS, M. (1979). *Class Control and Behaviour Problems*. Maidenhead: McGraw Hill, pp. xi+228, p. £3.95.
- S.G.B.F. (1979). *Bildungsforschung und Bildungspraxis*. Zug, Switzerland: Klett and Balmer Verlag for Schweizerische Gesellschaft für Bildungsforschung.
- SHARMAN, K. (1979). *Children's Television Behaviour*. Victoria: ACER, pp. 84, p. \$5.00.
- SKEMP, R. R. (1979). *Intelligence, Learning and Action*. Chichester: John Wiley and Sons, pp. x+324, c. £14.50.
- SONSTEGARD, M., SHUCK, A., and BEATTIE, N. R. (1979). *Living In Harmony With Our Children*. Luton: Millford Reprographics International, pp. 45, p. £1.00.
- STEINAKER, N. W., and BELL, M. R. (1979). *The Experiential Taxonomy—A New Approach to Teaching*. New York: Academic Press, pp. xvi+198, c. \$15.00.
- STERN, P. C. (1979). *Evaluating Social Science Research*. Oxford: Oxford University Press, pp. v+240, p. £3.50.
- STEWART, R. A., POWELL, G. E., and CHETWYND, S. J. (1979). *Person Perception and Stereotyping*. Farnborough: Saxon House, pp. ix+318, c. £9.50.
- STEWART, W. (1979). *Health Service Counselling*. Tunbridge Wells: Pitman Medical, pp. xi+175, p. £4.95.
- STURMAN, A. (1979). *From School to Work*. Victoria: ACER, pp. 7+114, p. \$5.00.
- STUART, S. (1979). *New Phoenix Wings. Preparation in Literature*. Henley on Thames: Routledge and Kegan Paul, pp. viii+189, c. £8.95.
- SUTTON, C. (1979). *Psychology for Social Workers and Counsellors—An Introduction*. Henley on Thames: Routledge and Kegan Paul, pp. xiv+230, c. £7.95, p. £4.95.
- THACKRAY, D. (Ed.) (1979). *Growth in Reading*. London: Ward Lock, pp. 282, n.p.
- TRAVERS, J. F. (1979). *Educational Psychology*. London: Harper and Row, pp. vii+740, p. £8.40.
- WALKER HEPNER, H. (1979). *Psychology Applied to Life and Work* (6th Ed.). Hemel Hempstead: Prentice-Hall, pp. xii+402, c. £10.35.
- WALLER, T. G., and MCKINNON, G. E. (Eds.) (1979). *Reading Research, Advances in Theory and Practice, Volume 1*. New York: Academic Press, pp. x+262, c. \$18.50.
- WALLER, T. G., and MCKINNON, G. E. (1979). *Reading Research, Volume 1*. London: Academic Press, pp. v+262, c. £12.00.
- WARD, C. (1979). *The Child in the City*. London: Penguin Books, pp. vi+221, p. £2.95.
- WARTELLA, E. (Ed.) (1979). *Children Communicating: Media and Development of Thought, Speech, Understanding*. London: SAGE, pp. 286, p. £5.00.
- WASS, H. (Ed.) (1979). *Dying: Facing the Facts*. Washington: Hemisphere, pp. xvi+426, p. \$10.95, c. \$19.95.
- WEIZMANN, R., BROWN, R., LEVINSON, P. J., and TAYLOR, P. A. (Eds.) (1979), *Piagetian Theory and the Helping Professions* (7th Annual Conference, Vol. I). California: University Affiliated Program, University of S. California, pp. 432, n.p.
- WERNER, E. E. (1979). *Cross-Cultural Child Development—A View from the Planet Earth*. California: Brooks/Cole, pp. xii+355, p. \$10.95.
- WICKER, A. W. (1979). *An Introduction to Ecological Psychology*. California: Brooks/Cole, pp. xii+228, p. \$6.95.
- WRIGHT, H. J., and PAYNE, A. N. (1979). *Evaluation of a School Psychological Service—The Portsmouth Pattern*. Winchester: Hampshire Education Department, pp. 69, n.p.
- WYLIB, R. C. (1979). *The Self Concept*. London: American University Pubs., pp. xiv+825, c. £21.00.
- YOUNGMAN, M. B. (1979). *Analysing Social and Educational Research Data*. London: McGraw-Hill, pp. v+235, p. £6.95.

Beyond Piaget:

Recent theories of concept development and their significance for teaching.

by Peter Langford

Published by the Centre for the Study of Urban Education at La Trobe University, Bundoora 3083, Victoria, Australia.

2nd Edition, 1979, 110 pages, illustrated.

Price: \$6 Aust., \$7 US, £3.50 sterling (postage included).

This book provides a systematic survey of post-Piagetian theories of concept development at an introductory undergraduate level, together with a discussion of their educational significance. The theories covered are Piaget, Klahr and Wallace's information processing theory, the 'dialectical' theory of Klaus Riegel and the author's own 'means of discovery' theory. This second edition also contains an appendix on Graeme Halford's neo-Piagetian structural theory.

LT008

The British Journal of Social and Clinical Psychology

Volume 18, Part 4, November 1979

Selected contents

- D. Miel, S. Duck & J. La Gaipa. Interactive effects of sex and timing in self-disclosure
A. J. Lazzarini, T. Cox & C. J. Mackay. Perceptions of and reactions to stressful situations
R. J. Brown & J. C. Turner. The Criss-Cross Categorization Effect in Intergroup discrimination
R. J. Edelmann & S. E. Hampson. Changes in non-verbal behaviour during embarrassment
J. L. Matson & R. A. Zeiss. The buddy system
E. Straube, U. Barth & B. König. Do schizophrenics use linguistic rules in speech recall?
D. G. Byrne. Anxiety as state and trait following survived myocardial infarction
D. Elton, G. D. Burrows & G. V. Stanley. The relationship between psychophysical and perceptual variables and chronic pain
T. N. Robinson & T. P. Zahn. Covariance of two-flash threshold and autonomic arousal for high and low scorers on a measure of psychoticism

Book reviews

Volume 19 (1980) £33.00 (US\$74.00). Single parts £10.00 (US\$22.00)

Volume 18 (1979) £25.00 (US\$57.00)

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN, UK

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

available from the Society's offices

St Andrews House, 48 Princess Road East, Leicester, LE1 7DR

Careers in Psychology	50p
Postgraduate Qualifications and Courses in Psychology	40p
Postgraduate Careers in Clinical Psychology	15p
Postgraduate Careers in Educational Psychology	15p
Postgraduate Careers in Occupational Psychology	15p
Postgraduate Careers in The Prison Service	15p
Postgraduate Careers in Social Psychology	15p

PRESIDENTIAL ADDRESSES

1968 Dissonance and Stress in the Teaching of Psychology: Some Reflections of an Accidental Psychologist: Professor George Seth	10p
1969 Changing Horizons: An Essay in Autobiography: Dr Boris Semenov	10p
1970 Choosing: Professor R. J. Audley	10p
1971 In Search of an Idiographic Psychology: Professor H. Gwynne Jones	10p
1972 Psychology Today and Tomorrow: Professor Harry Kay	10p
1973 Psychology in Society: Ends or End: Professor Max Hamilton	10p
1974 On Taking Sides: Professor B. M. Foss	10p
1975 Thought and the Brain: Professor O. L. Zangwill	10p
1976 Psychology and Social Policy: Professor Jack Tizard	10p
1977 The Scientific/Applied Debate in Psychology: A Contribution: May A. Davidson	20p
1978 Predicting Human Development: Professor A. D. B. Clarke	20p

C. S. MYERS LECTURES

1964 Remembering Dr Myers: Professor F. C. Bartlett	10p
1965 Personality and Experimental Psychology: Professor H. J. Eysenck	10p
1967 Practice a liking: A study in experimental aesthetics: Professor D. W. Harding	10p
1968 What is potential ability?: Professor P. E. Vernon	10p
1970 In defence of empirical psychology: Dr E. E. Broadbent	10p
1971 C. S. Myers in retrospect: Professor Alec Rodger	10p
1973 On applied psychology in Britain: Professor George Drew	10p
1975 Measuring psychiatric morbidity in a South Indian population: Dr G. M. Carstairs	10p
1976 Quantitative subjective assessments are almost always biased: Sometimes completely misleading: Dr E. C. Poulton	10p
1977 The neuropsychology of anxiety: Dr J. A. Gray	10p

REPORTS TO GOVERNMENT DEPARTMENTS

To the Royal Commission on the Law Relating to Mental Illness and Mental Deficiency 1954-1957	10p
To the Royal Commission on the Penal System in England and Wales	10p
To the Royal Commission on Medical Education	10p
To the Ministry of Health and Department of Education and Science on the Responsibility for the Education of Mentally Handicapped Children	10p
To the Advisory Council on the Penal System (Sub-committee on Detention Centres)	10p
To the Departmental Committee on the Adoption of Children	10p
To the Department of Health and Social Security Sub-committee on the Role of Psychologists in the Health Services	10p
To the Butler Committee on the Law relating to the Mentally Abnormal Offender	10p
Comments on the DHSS Review of the Mental Health Act 1959: Professional Affairs Board: <i>Bulletin</i> (1977), 30, 223-226	10p
Evidence to the Data Protection Committee (Home Office): Professional Affairs Board: <i>Bulletin</i> (1977), 30, 224-248	10p
Comments on the Report of the Local Committee on Child Health Services: Professional Affairs Board: <i>Bulletin</i> (1978), 31, 157-160	10p
Comments on the DES draft circular on the Discovery of Handicapped Pupils and the Assessment of their Needs for Special Education: Professional Affairs Board: September 1974	10p

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

(continued)

REPORTS TO GOVERNMENT DEPARTMENTS—continued

Evidence to the Royal Commission on the National Health Service: Professional Affairs Board: March 1977	10p
Comments on the Draft Document on Prevention in the Child Health Services: Professional Affairs Board: May 1978	10p
Evidence to the Home Office Review Committee on the Prison Psychological Service: Professional Affairs Board: May 1978	10p
Comments on the draft health circular on Health Centres and Other Provision for Primary Health Care: Professional Affairs Board: May 1978	10p
Comments on the DHSS Document on Health Service Complaints Procedure: Professional Affairs Board: August 1975	25p
Comments on "A Happier Old Age": A discussion document from the DHSS and Welsh Office: Professional Affairs Board: September 1978	25p
Manpower Survey: Psychologists and Mental Handicap: Professional Affairs Board: November 1978	25p
Comments on N.D.G. Pamphlet No. 5: Professional Affairs Board: November 1978	25p
Evidence to the Royal Commission on Criminal Procedure: Professional Affairs Board: November 1978	25p
Response to the Consultative Document on the Warnock Report: Professional Affairs Board: March 1979	25p
Evidence to the DHSS Working Party on Observation and Assessment: Professional Affairs Board: 1979	25p

OTHER REPORTS

Principles governing the Employment of Psychological Tests and Clinical Instruments	7p
Psychological Tests: A Statement by The British Psychological Society	7p
Teaching Educational Psychology in Training Colleges	25p
Technical Recommendations for Psychological and Educational Tests	10p
The Psychological Study of Deafness and Hearing Impairment: Edited by M. Rodda	40p
The changing role of the Educational Psychologist in Child Guidance Services in Scotland	10p
Teaching Psychology to Nurses (DCP Report)	10p
Training in Psychotherapy (DCP Report)	10p
Guidelines for the In-Service Training of Probationer Clinical Psychologists	10p
The Work of the Conference Programme by Dr George Delafield	10p
The Teaching of Psychology in Schools: Joint Statement by The British Psychological Society and the Association for the Teaching of Psychology: September 1977	15p
British Psychological Education: Scientific Affairs Board: November 1977	20p
Ethical Principles for Research with Human Subjects: Scientific Affairs Board: April 1978	15p
Comments on the Proposed New CCETSW Training Programme: Professional Affairs Board: May 1978	25p
Report of the Working Party on the Sex Discrimination Act: Professional Affairs Board: May 1978	25p
Report of the Working Party on Behaviour Modification: Professional Affairs Board: September 1978	25p
Report of the Working Party on Abuses of Psychology for Political Purposes: Professional Affairs Board: October 1978	25p
Notes for guidance in planning short courses on Psychological Testing: Standing Committee on Test Standards: February 1979	25p
Postgraduate training in Clinical Psychology: Optimal length and format of training: Professional Affairs Board: <i>Bulletin</i> (1977), 30, 284-285	25p
Computer Use in British Psychology Departments: David Hale or Scientific Affairs Board: <i>Bulletin</i> (1977), 30, 410-413	25p
Repertory Grid Methods: Maths and Stats Psychology Section	50p
Monitoring Financial Support for Psychological Research: Scientific Affairs Board: <i>Bulletin</i> (1978), 31, 3-5	25p
Psychological Services for Children: Professional Affairs Board: <i>Bulletin</i> (1978), 31, 125-127	15p
Report of Working Party on Animal Experimentation: Scientific Affairs Board: <i>Bulletin</i> (1979), 32, 44-52	25p

International Journal of PSYCHOLOGY

Editor

(Scientific Papers):

G. d'YDEWALLE
Univ. of Leuven, Belgium

Associate Editor

(International Platform):

W. HOLTZMAN
Univ. of Texas, Austin,
U.S.A.

Editorial Board:

a) The IUPS Executive Committee -

A. SUMMERFIELD
England
B. LOMOV
USSR
R. DIAZ-GUERRERO
Mexico
D. BELANGER
Canada
W. HOLTZMAN
U.S.A.
K. PAWLIK
W. Germany
M. O. A. DUROJAIYE
Nigeria
M. FRANKENHAEUSER
Sweden
R. FRAISSE
France
J. GUEVARA
Cuba
F. KLIX
Germany
J. NUTTIN
Belgium
M. ROSENZWEIG
U.S.A.
R. RUSSELL
Australia
T. TOMASZEWSKI
Poland

b) Consulting Editors

R. ARDILA
Colombia
M. COLE
U.S.A.
A. GINSBERG
Brazil
B. S. GUPTA
India
G. JAHODA
Scotland
M. YELA
Spain

The INTERNATIONAL JOURNAL OF PSYCHOLOGY is edited by the International Union of Psychological Science (IUPS). Its purpose is to facilitate communication, to develop understanding, and to foster cooperation among psychologists throughout the world.

The main part of each issue is devoted to scientific and theoretical papers in all fields of psychology, with emphasis on topics where the social or cultural context is important; fields covered will include general psychology (perception, learning, cognitive processes, language, etc.), developmental psychology, personality, and social psychology. The Journal emphasizes basic research and theory rather than technical and applied problems.

Many issues of the Journal also include a second part, the International Platform for Psychologists, which provides an opportunity to exchange news and opinions on psychology as an academic and applied profession.

Beginning with Volume 14, 1979, the publication and distribution of the INTERNATIONAL JOURNAL OF PSYCHOLOGY/JOURNAL INTERNATIONAL DE PSYCHOLOGIE are handled by North-Holland Publishing Company, Amsterdam, The Netherlands.

Publication Schedule:

1979 Volume 14 in 4 issues.

Subscription Information:

Institutional Subscriptions:

The subscription price for Volume 14 (1979) is US \$44.00/Dfl. 90.00 including postage and handling.

Personal Subscriptions:

Private subscribers are entitled to a subscription at the reduced rate of US \$27.00/Dfl. 55.00 including postage and handling.

Personal subscriptions must be prepaid; the orders must be placed directly with the Publisher and copies should not be made available to institutions.

The Dutch guilder price is definitive. US \$ prices are subject to exchange rate fluctuations.

Orders and requests for specimen copies may be sent to the Publisher:

north-holland

P.O. BOX 211
1000 AE AMSTERDAM
THE NETHERLANDS

IN THE U.S.A. AND CANADA:
52 VANDERBILT AVENUE
NEW YORK, N.Y. 10017

The Dutch guilder price is definitive, US \$ prices are subject to exchange rate fluctuations

6030 NHa

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
Journal of Occupational Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1980 subscription prices, payable in advance, are £35.00 (US \$77.00 in the USA and Canada) for *The British Journal of Psychology*; £33.00 (US \$74.00) for *The British Journal of Social and Clinical Psychology*; £28.00 (US \$62.00) for *The British Journal of Medical Psychology* and £25.00 (US \$55.00) for *Journal of Occupational Psychology*.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £25.00 per volume post free (US \$55.00 in the USA).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK. A special price of £130 (\$290) for all five journals going to the same address is offered for 1980.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £12.00 a year, post free, or £5.00 each number. The prepaid subscription price in the USA is \$25.00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

MANUSCRIPTS

For *The British Journal of Psychology* should be sent to Professor M. Coltheart, Department of Psychology, Birkbeck College, Malet Street, London WC1E 7HX.

For *The British Journal of Medical Psychology* should be sent to Professor J. P. Watson, Department of Psychiatry, Guy's Hospital Medical School, London Bridge, London, SE1.

For *Journal of Occupational Psychology* should be sent to Professor P. Warr, MRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For *The British Journal of Social and Clinical Psychology* should be sent to G. M. Stephenson (Social Psychology Editor), Professor of Social Psychology, Eliot College, University, Canterbury, Kent, CT2 7NS; or to Dr. H. R. Beech (Clinical Psychology Editor), Department of Psychiatry, Withington Hospital, West Didsbury, Manchester, M20 8LR;

For *The British Journal of Mathematical and Statistical Psychology* should be sent to Professor P. Levy, Department of Psychology, University of Lancaster, Lancaster, LA1 4YF;

For *The British Journal of Educational Psychology* should be sent to Professor Dennis Child, School of Education, St. Thomas' Street, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2.00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £20.00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £4.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

23 SEP 1981 VOL. 50, PART 1, 1980

CONTENTS

EDITORIAL 1

ARTICLES:

- | | | |
|---|--|----|
| PHINNEY, J. S. and
FESHBACH, N. D. | Non-directive and Intrusive Teaching Styles of Middle-
and Working-class English Mothers | 2 |
| SILVA, P. A. and
BRADSHAW, J. | Some Factors Contributing to Intelligence at Age of
School Entry | 10 |
| HICKS, C. | The ITPA Visual Sequential Memory Task; an Alter-
native Interpretation and the Implications for Good and
Poor Readers | 16 |
| DOWLING, J. R. | Adjustment from Primary to Secondary School: a One
Year Follow-up | 26 |
| FRY, P. S. and COB, K. J. | Interaction Amongst Dimensions of Academic Motiva-
tion and Classroom Social Climate: a Study of the
Perceptions of Junior High and High School Pupils | 33 |
| YOUNGMAN, M. B. | Some Determinants of Early Secondary School Perform-
ance | 43 |
| BAR-TAL, D., KFIR, D.,
BAR-ZOHAR, Y. and
CHEN, M. | The Relationship between Locus of Control and Academic
Achievement, Anxiety and Level of Aspiration | 53 |
| BEDWELL, C. H., GRANT, R.
and McKEOWN, J. R. | Visual and Ocular Control Anomalies in Relation to
Reading Difficulty | 61 |

RESEARCH NOTES:

- | | | |
|----------------------------------|---|----|
| MAQSUD, M. | Extraversion, Neuroticism, Intelligence and Academic
Achievement in Northern Nigeria | 71 |
| HARVEY, T. J. and
EDWARDS, P. | Children's Expectations and Realisations of Science | 74 |
| CONFERENCE ABSTRACTS | | 77 |
| PUBLICATIONS RECEIVED | | 94 |

THE BRITISH
JOURNAL OF
EDUCATIONAL
PSYCHOLOGY

JUNE 1980

VOL. 50, PART 2

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by
Professor DENNIS CHILD

Assistant Editors: Sub-Editor:
Professor S. N. BENNETT Dr. S. CUNNINGHAM
Dr. G. BROWN
Professor K. WEDELL

Advised by the Editorial Committee:

Dr. H. BELOFF	Professor E. A. PEEL
Professor R. A. KENNEDY	Mr. P. PRINCEP
Professor P. LEVY	Professor B. SHACKEL
Professor J. D. NISBET	Dr. D. SHARPLES
Professor M. SUTHERLAND	

Vol. 50, Part 2

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £12.00 a year post free, payable in advance, or £5.00 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication should be sent to Professor Dennis Child, School of Education, St. Thomas' Street, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All Books for Review for the *British Journal of Educational Psychology* should be sent to: Dr. G. Brown, Department of Educational Research, The University of Lancaster, Bailrigg, Lancaster.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by
SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1980. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY R. & R. CLARK LTD., EDINBURGH

THE INFANT RATING SCALE

By G. A. LINDSAY

(Psychological Service, Sheffield Education Department)

SUMMARY. A study of the Infant Rating Scale (IRS) was made to examine its usefulness as an instrument for the early identification of learning difficulties. 1,341 children, divided into three samples by terms of entry, were rated by their teachers after one term in school. The structure of the IRS, its reliability and predictive validity are examined. It is concluded that the IRS is a useful aid to the teacher of young children in the early identification of learning problems.

INTRODUCTION

A MAJOR concern over the past decade has been the early identification of children with learning difficulties, or 'screening' as it is commonly known (e.g., Bullock Report, DES, 1975). The focus of this work has changed from children with significant problems at the time of investigation, whether they be babies with severe intellectual handicaps, or eight-year-olds with reading problems, to attempts to identify children before the problems became manifest.

Hence the emphasis has shifted to attempts to identify children 'at risk' at a pre-symptomatic stage. As a consequence, the measures used have altered from variables such as reading to supposed prerequisite skills and attributes such as visual perception and use of language.

Another change in emphasis has been away from the use of tests (e.g., De Hirsch *et al.*, 1966) to the use of teacher rating scales, (e.g., Myklebust, 1971; Novack *et al.*, 1973). The reasons for this range from the financial to the practical (the problems of testing large numbers of young children). On such scales the class teacher is asked to judge the child on a number of dimensions usually including cognitive and behavioural functioning. Typically ratings may be on a 1-5 scale ranging from high competence to deficits or problem behaviour, or of a Yes/No type indicating the presence or absence of competencies or worrying aspects of development. There is evidence that the teacher rating scales are as useful as tests at identifying children who will have learning difficulties, and indeed that they can be more accurate (Feshbach *et al.*, 1974).

A major problem, however, is that many of the instruments available have a limited amount of evidence for their usefulness which renders evaluation difficult. This is particularly so for the best-known English instruments, the Croydon Checklist (Wolfendale, 1976) and the Screening Instrument (Rawsthorne, 1967; Tansley, 1976). Indeed the latter is now used extensively in one major English city although "no large scale statistical enquiry has been carried out to measure the progress of the children concerned" (Gray and Reeve, 1978). The only study which evaluated this instrument appears to be that of Lindsay (1974) which failed to show that it was a useful device.

Consequently, when the Sheffield LEA examined this issue of identifying children with learning difficulties it was decided to carry out a research project which not only produced an instrument but also examined its effectiveness.

METHOD

Procedure

The total intake of new entrants into 21 infant and first schools, chosen to represent the variety of population in the city, apart from schools with large numbers of children of Asian origin, was assessed by their class teachers using the Infant Rating Scale (IRS) after they had been in school for 1-1½ terms. In the second term

of the final infant year, one-third of the total sample was given a reading test. In the final term, the total sample was given the Top Infant Edition of the Infant Rating Scale (IRS-TOP).

Sample

A total of 1,341 children (679 girls, 662 boys; mean age 62.69 months, sd 1.43 months) were administered the IRS. As each term's intake was rated separately, the results from three cohorts are available. The sample had reduced to 916 by the end of infant school when the IRS-TOP was administered. While most of this reduction was a result of children changing school, part was caused by some teachers being unable to rate all their children owing to time constraints. In that case they were asked to rate 50 per cent randomly.

All the sample children attending seven of the schools were given the Young Group Reading Test ($N = 480$). These schools were chosen to reflect the schools in the total sample.

Measures

The IRS comprises 25 items each of which is rated on a 1-5 scale according to a booklet of instructions with point 5 representing the 'problem' end. The items were originally grouped *a priori* under six headings; the groupings were Behaviour, Expressive Language, Receptive Language, Motor Skills, Early Learning Abilities, and General Development. Table 1 lists the items (see Lindsay, 1979b).

The IRS was designed to reflect a distribution of scores such that about 1-2 per cent of children would be rated at point 5 for each item, and about 10-13 per cent would be rated at point 4, so giving an overall proportion of about 15 per cent of

TABLE 1
PERCENTAGE OF CHILDREN RATED ON EACH IRS VARIABLE ($N = 1341$)

	1	2	Rating 3	4	5
Temperament	17.9	41.4	31.3	8.2	1.2
Attitude	24.8	39.9	26.4	8.3	0.6
Kindness	12.6	44.6	37.3	5.1	0.4
Concentration	13.5	34.7	35.9	13.3	2.6
Participation	10.4	50.0	32.5	6.2	1.0
Acceptance	6.0	44.5	44.6	5.0	0.0
Desire to mix	6.8	38.7	50.4	3.3	0.8
Response to new situations	6.3	27.4	50.4	14.5	1.5
Articulation	11.0	34.2	42.0	10.0	2.8
Vocabulary	4.5	25.6	55.8	12.2	2.0
Sentence construction	6.5	29.2	51.5	11.0	1.9
Expression	7.2	36.6	44.9	8.9	2.4
Comprehension of instructions	6.8	36.4	44.5	10.9	1.4
Understanding of words	6.3	31.5	56.3	5.1	0.7
Memory	5.1	31.8	52.5	9.5	1.0
Gross motor	1.9	19.4	70.6	7.2	0.8
Fine motor	8.2	26.0	49.7	14.1	2.0
Drawing	7.5	35.1	42.4	12.1	2.9
Writing	3.8	28.7	54.6	11.0	1.8
Reading	29.5	31.5	30.3	8.3	0.3
Number	3.5	29.5	55.5	9.6	1.8
Approach to learning	11.9	28.8	46.5	11.8	1.0
Attention	6.8	27.3	40.5	21.8	3.5
Maturity	16.5	39.3	33.2	9.8	1.1
Level of concern felt	11.1	32.4	43.7	10.8	2.0

children 'at risk' as suggested by studies such as Davie *et al.* (1972), Rutter *et al.* (1970). It was not expected that a normal distribution of scores need obtain.

The IRS-TOP was of similar format, but with some new items (e.g., spelling) and with other items modified, where necessary, to reflect seven-year-old children's development (Lindsay, 1979c).

On each scale, a score was available for each item and, by summation, a total score could be computed.

The Young Group Reading Test (Young, 1968) was administered according to the manual.

RESULTS

Structure of the IRS

(i) *Distribution of scores.* The overall distribution of scores on each item is presented in Table 1. The percentage of children rated at point 5 varies from 0 per cent (Acceptance by Peers) to 3.5 per cent (Attention and Distractibility). The percentage of children rated at point 4 varies from 3.3 per cent (Desire to Mix) to 21.8 per cent (Attention and Distractibility). Hence the anticipated distribution was not found for each variable. However, the respective percentages of children for the general variable Level of Concern Felt were 2.0 per cent (point 5) and 10.7 per cent (point 4).

There were highly significant sex differences in mean scores ($P < 0.001$) on 20 of the items, with girls always being rated better than boys. Three items (Participation, Desire to Mix and Gross Motor Skills) were non-significant, while Vocabulary and Understanding of Words were significant at the 0.01 level.

(ii) *Factor analysis.* In order to examine the internal structure of the IRS more fully, the results of each term's sample were factor analysed. The methods chosen were principal components with iterations, followed by a varimax rotation of the significant principal components. The Statistical Package for the Social Sciences (Nie *et al.*, 1975) was used for the analyses.

The results of the principal components analysis for each term's sample are presented in Table 2. A close similarity is evident across each sample. In each case about 50 per cent of the variance is accounted for by one principal component.

TABLE 2
RESULTS OF PRINCIPAL COMPONENTS ANALYSIS OF THE IRS FOR EACH TERM'S SAMPLE

Factor	Eigen Value			Percentage of Variance		
	Term 1	Term 2	Term 3	Term 1	Term 2	Term 3
1	12.654	13.197	13.404	50.6	52.8	53.6
2	2.118	2.226	2.104	8.5	8.9	8.4
3	1.656	1.289	1.320	6.6	5.2	5.3
4	1.081	0.937	1.118	4.3	3.7	4.5
5	0.791	0.888	0.832	3.2	3.6	3.3

The general principle of rotating principal components with eigenvalues greater than unity was followed but as there were only three such components for the term 2 sample, a different factor structure was produced. It was decided, therefore, to re-analyse, reducing the criterion for the minimum eigenvalue to produce four rotated factors.

Table 3 presents the significant factor loadings (> 0.50) for each variable. The

TABLE 3
MATRIX OF SIGNIFICANT ROTATED FACTOR LOADINGS OF THE INFANT RATING SCALE

Variable	Factor 1			Factor 2			Factor 3			Factor 4		
	Term			Term			Term			Term		
	1	2	3	1	2*	3	1	2*	3	1	2	3
Temperament				64	66	67						
Attitude				78	79	81						
Kindness				71	73	71						
Concentration				55				59				
Participation										70	68	71
Acceptance										75	73	80
Desire to mix										81	82	83
Response to new situations												
Articulation	72	72	72									
Vocabulary	82	86	84									
Sentence construction	83	84	77									
Expression	79	74	77									
Comprehension of words	53	52	61					53				
Understanding	69	72	73									
Memory	67	64	69									
Gross motor skills							70	54	73			
Fine motor skills							72	61	69			
Drawing							64	67	71			
Writing								58				
Reading								65	51			
Number			50									
Approach to learning				58		52		69				
Attention				69		53		70	51			
Maturity							53	52				
Level of concern felt		50	51				53	54				

Decimal points omitted.

* N.B.—For term 2, Factors 2 and 3 have been transposed.

results of the first and third term's samples show great similarity. Factor 1 is clearly a language factor. Factor 2 is a behavioural factor, with loadings on the 'temperamental' and 'approach to learning' items. Attention and Concentration also tend to load significantly. The third factor relates mainly to fine motor skills, writing and drawing, with the other learning variables loading at or just below the 0.50 level. Finally, Factor 4 is a second behavioural factor, but relating to the child's relationships with his peers.

The variables 'Level of concern felt' and 'Maturity' tend to load moderately on more than one factor, which reflects their relating to the general development of the child. A similar situation applies with 'Response to new situations' though the loadings are lower, and none reaches significance. The variable 'Gross motor skills' does not load significantly on any factor.

When the results of the second term's sample were re-analysed, using a minimum eigenvalue of 0.9 to govern selection of principal components for rotation, a similar result to the analysis of other terms' results was obtained, except that the second and third factors were transposed. For the sake of clarity, however, these two factors are transposed in Table 3 in line with the structure of terms 1 and 3.

Because of the close similarity across each term's data, an analysis of the total sample produced results which closely reflect those presented here.

Reliability

The only measure of reliability which was considered relevant to assess was stability over time. While inter-rater reliability would provide another useful source of information, this was not possible to assess owing to the usual practice of each class having only one teacher who knew them well. It was not considered appropriate for them to be rated by someone less familiar with them, e.g., headteacher.

A total of 49 children were rated twice on the IRS with a two-week gap between each rating. On each occasion the child was rated by his class teacher. The ratings were carried out by 17 teachers in five schools. (The forms for the first rating were removed by the headteacher immediately after completion.)

Table 4 shows the high level of stability for each variable over this short time interval. Test-retest correlations range from 0.65 to 0.93 for individual items (median 0.85) while the IRS Total Score shows a reliability coefficient of 0.96.

TABLE 4

PEARSON CORRELATION COEFFICIENTS FOR TEST-RETEST VALUES OF EACH IRS VARIABLE (N = 49)

Variable	r
Temperament	0.84
Attitude to teacher	0.83
Kindness to peers	0.79
Concentration and ability to organise	0.88
Participation in class activity	0.71
Social integration—acceptance	0.77
Social integration—desire to mix	0.65
Response to new situations	0.82
Articulation	0.93
Vocabulary	0.92
Sentence construction	0.92
Expression	0.86
Comprehension of instruction	0.87
Understanding of words	0.89
Memory for oral information	0.86
Gross motor skills	0.79
Fine motor skills	0.82
Drawing	0.84
Writing	0.89
Reading	0.89
Number	0.89
Approach to learning	0.85
Distractibility	0.86
Maturity	0.80
Level of concern felt	0.89
Total IRS Score	0.96

P < 0.001 in every case

It is interesting that the variables showing least stability relate to social behaviour (e.g., Desire to mix 0.65, Participation in class activity 0.71), whilst those showing greatest stability relate to cognitive skills, particularly language (e.g., Vocabulary 0.92, Sentence Construction 0.92).

Predictive validity

The predictive validity of the IRS was measured against the Young Group Reading Test and the IRS-TOP administered in the second and third terms respectively of the final infant year. The time span, therefore, was 2-2½ years. The data were examined using both a correlation method (Pearson's product-moment) and probability statistics. Only the former results are presented here.

Factor scores for the IRS were computed by summing the ratings of the items which loaded significantly (>0.50) on each factor, as derived from the factor analysis. These factor scores, and the total score, were correlated against the Young Reading Test. These results are presented in Table 5.

TABLE 5
PEARSON CORRELATION COEFFICIENTS OF YOUNG READING TEST (STANDARD SCORE) AGAINST IRS (AS 4 DERIVED FACTORS) (N = 480)

IRS factor	r
Behaviour	-0.24
Social integration	-0.22
Language	-0.48
Aspects of learning	-0.48
Total Score	-0.45

P < 0.001 in each case

It will be seen that there is a difference in predictive power between the cognitive factors (Language, Aspects of Learning) and the behavioural factors (Social Integration, Behaviour) although coefficients were highly significant ($P < 0.001$).

The correlation coefficients of the IRS items against their comparable IRS-TOP items are presented in Table 6. Here the correlation of total scores was 0.61, with a range for individual items of 0.26 (Gross motor skills) to 0.54 (Level of concern felt). In general, the cognitive items again showed higher levels of correlation than the behavioural items.

DISCUSSION

The results presented here suggest that IRS may be considered a useful instrument for the teacher of young children. As such the data complement those derived from a study using a sub-sample of the standardisation sample which showed the IRS to be a better predictor of later achievement (at seven years) than a variety of individually administered tests (Lindsay, 1980).

The IRS is reliable over a short period of time; its factor structure is coherent and consistent over three independent samples; and the correlation of 0.61 of the total score with the IRS-TOP total score is reasonable. However, the correlation coefficients of the factor scores with later reading though highly significant, are not high (up to 0.48) and not all items have distributions of ratings close to those projected.

Some modifications have been shown to be necessary. The original six groupings of items were not substantiated by the factor analysis, where only four factors were found. The item 'Gross motor skills' was found not to be of use when both the construct and predictive properties of the scale were examined. Also, the Reading

TABLE 6

PEARSON CORRELATION COEFFICIENTS FOR THE
IRS-TOP EQUIVALENT ITEMS (N = 916)

IRS items	IRS-TOP correlation coefficients (equivalent items)
Temperament	0.32
Attitude to teacher	0.43
Kindness to peers	0.39
Participation in class activity	0.23
Acceptance by peers	0.35
Desire to mix	0.29
Articulation	0.48
Vocabulary	0.45
Sentence construction	0.44
Expression	0.48
Comprehension	0.44
Word understanding	0.44
Memory	0.43
Concentration	0.51
Approach to learning	0.48
Attention and distractibility	0.53
Fine motor	0.38
Drawing	0.41
Writing	0.45
Reading	0.46
Number	0.42
Gross motor	0.26
Response to new situations	0.37
Concern felt	0.54
Total Score correlation	0.61

P < 0.001 in all cases

item was not considered to show a suitable distribution of scores. The present edition of the IRS takes note of these points (Lindsay, 1979b).

While the IRS can, on this evidence, be recommended for use in the early identification of learning difficulties, it is necessary to note that further evidence from this study (Lindsay, 1979d) has raised questions about the usefulness of the 'at risk' concept. These arguments will be extended in a subsequent paper.

In the light of these points it is recommended that the IRS be seen as a useful instrument which helps structure teachers' observations, and gives pointers to children's difficulties, but that this process take place within the context of a continuous monitoring system with in-built intervention (Lindsay, 1979a) rather than a one-off screening for the purpose of classifying children as at risk of educational failure.

ACKNOWLEDGMENTS.—I am grateful to the children and teachers involved, and to Sheffield LEA for enabling the project to take place.

REFERENCES

- DAVIE, R., BUTLER, N., and GOLDSTEIN, H. (1972). *From Birth to Seven*. London: Longmans.
 DE HIRSCH, K., JANSKY, J. J., and LANGFORD, W. S. (1966). *Predicting Reading Failure*. London: Harper and Row.
 D.E.S. (1975). *A Language for Life*. (The Bullock Report.) London: H.M.S.O.
 FESHBACH, S., ADELMAN, H., and FULLER, W. (1974). Early identification of children with high risk of reading failure. *J. learn. Dis.*, 7, 639-644.
 GRAY, D., and REEVE, J. (1978). Ordinary schools: some special help. *Special Educ.*, 5, 25-27.

- LINDSAY, G. A. (1974). Early identification of learning difficulties. Unpublished M.Ed. (Ed.Psych.) Dissertation, University of Birmingham.
- LINDSAY, G. A. (1979a). The early identification of learning difficulties and the monitoring of children's progress. Unpublished Ph.D. Thesis, University of Birmingham.
- LINDSAY, G. A. (1979b). *The Infant Rating Scale: Manual*. Sheffield Education Department. (In press.)
- LINDSAY, G. A. (1979c). *The Infant Rating Scale—Top Infant Edition*. Sheffield Education Department. (In press.)
- LINDSAY, G. A. (1979d). The early identification of learning difficulties: a critique. Paper presented to the symposium on Screening, Education Section of the British Psychological Society, Oxford.
- LINDSAY, G. A. (1980). The Infant Rating Scale: some evidence on its validity from a selected sample of children. *Division of Educational and Child Psychology Occasional Papers*, 3 (2).
- MYKLEBUST, H. R. (1971). *The Pupil Rating Scale*. New York: Grune and Stratton.
- NIE, N., HULL, C., JENKINS, J., STEINBRENNER, K., and BENT, D. (1975). *Statistical Package for the Social Sciences*. (2nd edit.). New York: McGraw-Hill.
- NOVACK, H., BONAVENTURA, E., and MERENDA, P. (1973). A scale for the early identification of children with learning problems. *Except. Child.*, 40, 98-104.
- RAWSTHORNE, K. J. (1967). A procedure for screening children in their final year in the infant school. Unpublished Dip.Sp.Ed. (ESN) Thesis, University of Birmingham.
- RUTTER, M., TIZARD, J., and WHITMORE, K. (1970). *Education, Health and Behaviour*. London: Longmans.
- TANSLEY, A. E. (1976). Special educational screening in infant schools. In WEDELL, K., and RAYBOULD, E. C. (Eds.), *The Early Identification of Educationally 'at-risk' Children*. *Educ. Rev.*, Occasional Publication No. 6, University of Birmingham.
- WOLFENDALE, S. (1976). Screening and the early identification of reading and learning difficulties—description of the Croydon screening procedures. In WEDELL, K., and RAYBOULD, E. C. (Eds.), *The Early Identification of Educationally 'at-risk' Children*. *Educ. Rev.*, Occasional Publication No. 6, University of Birmingham.
- YOUNG, D. (1968). *Manual for the Young Group Reading Test*. London: University of London Press.

(Manuscript received 5th July, 1979)

THE COMPARATIVE EFFECTS OF ADULT AND PEER GROUP INFLUENCE ON CHILDREN'S MORAL JUDGMENTS

By M. SIEGAL

(Department of Psychology, University of Queensland)

AND M. C. BOYES

(Department of Psychology, University of British Columbia)

SUMMARY. This paper examines the comparative effects of adult and peer group influence on children's moral judgment development. Little support is found for Piaget's claim that with increasing peer group interaction the authority of the adult becomes less legitimate and influential to the child. The possibility is suggested that the adult exerts increasing influence with age on children's moral judgments, and that perhaps less emphasis should be placed on peer group centred moral education programmes and more stress put on the effectiveness of direct adult intervention.

INTRODUCTION

ONE of the most central and enduring issues in social and personality development concerns the comparative effects of adult and peer group influence on children's moral development. Pronouncements have been made on this theme ever since the time of Ancient Greece:

"The affection of children for their parents (like man's love of God) is the sort of feeling one has for what is good and superior. For their parents have bestowed on them the greatest of blessings—they have given them life and nursed them and provided for their education when they reached school age. Beyond this, family affection has more pleasure and profit in it than has any friendship between unrelated persons in proportion as family life provides a better chance of going shares in everything" (Aristotle (385-322 B.C.) as translated by Thomson (1953, p. 251)).

This question has also occupied the writings of contemporary developmental psychologists who have given it a very different twist by relating the comparative influence of adults and peers to stages of moral judgment development. To this end, peer group influence has often been assigned a prominent role in contributing to the process of stage transition. According to Piaget (1932/1977), there are generally two childhood moralities: one characterised by adult constraint and a unilateral respect for adult rules, the other characterised by a mutual respect, co-operation and solidarity among peers. Although elements of both types of moralities are present in older children and no precise ages can be attached to a stage transition, adult constraint is in general the morality of young children under 9 years. For the young child, moral rules are the infallible product of adult authority and should not be violated even accidentally or with good intentions. Adult constraint children judge actions mainly on the basis of consequences. They maintain that the well-intentioned child who accidentally causes heavy damage is naughtier than the ill-intentioned child who deliberately causes light damage.

In contrast, the older child is more apt to consider intentions when making a moral judgment. Moral rules are not imposed by an external authority but are the produce of lawful convention and co-operation among equals. Piaget does not deny that young children demonstrate some semblance of a mutual respect morality, only that older children display more mutual respect than do younger ones. This development and the circumstances under which young children do make intentional

moral judgments has been documented in subsequent studies (cf. Karniol, 1978; Keasey, 1977; Siegal, 1975).

One of the major catalysts for the transition to a morality of co-operation and mutual respect is an increasing peer group interaction which, according to Piaget, 'liberates' the child from adult constraint. With the advent of peer group influence, the authority of the adult becomes less legitimate to the child: "The main factor in the obligatory conformity of very young children is nothing but respect for age—respect for older children, and, above all, respect for adults. And if, at a given moment, co-operation takes the place of constraint, or autonomy that of conformity, it is because the child, as he grows older, becomes progressively free of adult supervision" (Piaget, 1932/1977, p. 99).

Piaget's ideas on the subject of moral development have sparked a voluminous amount of research. They have become the cornerstone for well-known theories which aim to describe moral development through adolescence and adulthood (e.g., Kohlberg, 1963, 1976, 1978). In this paper, we examine the research as it pertains to the comparative effects of adult and peer group influence. It will be suggested that, while there are some areas in which adult authority does wane with an increase in peer group interaction, adult authority may actually become more influential in the area of moral judgment. This interpretation will be discussed in terms of its educational implications.

STUDIES ON THE COMPARATIVE EFFECTS OF ADULT AND PEER INFLUENCE

Judgments of acts committed by adult and peer transgressors

A number of testable predictions can be derived from Piaget's adult constraint notion. For example, it implies that children, having a unilateral respect for adult authority, should judge adult transgressors less severely than peer transgressors. Yet several recent studies have indicated that young children's moral judgments are not affected in this way by the age of the transgressor. Rybash *et al.* (1975) presented kindergarteners (mean age = 6 years) with a single moral problem in which either an adult or a peer accidentally broke a number of Easter eggs. The adult and peer were judged as equally naughty, regardless of their age. Similar findings have been reported by Suls and Kalle (1978). Children aged 5 to 10 were given four moral problems involving, for example, an adult or a peer who broke all the dishes. Again adults were not judged less harshly than children causing identical damage, even by subjects as young as 5 years.

In both of these studies the children were asked to rate separately the behaviour of adult and peers. In a study by Peterson *et al.* (1974) children were required to make a direct comparison. Preschoolers (mean age = 4 years), second graders (mean age = 7 years), and adults were given story pairs which included a contrast between a well-intentioned act causing heavy damage with an ill-intentioned act causing light damage. Piaget's theory would predict that young children should be more likely to excuse a well-intentioned act causing heavy damage when the transgressor is an adult rather than a peer. Yet, consistent with the findings of Rybash *et al.* and Suls and Kalle, preschoolers as well as adults evaluated adult and peer transgressors as equally naughty. Only the second graders' judgments appeared to be consistent with Piaget's theory, and it is reported that this group evaluated adult transgressors more favourably than peers. However, these results may have been analysed inappropriately. Though the combined data came from two related samples, the analysis was done using chi-square tests for independent samples. Fortunately, half of the children received a different, equivalent set of stories. This procedure permits a re-analysis of data drawn from two independent samples: 12 (48%) of 25 second-graders judged a well-intentioned adult favourably as opposed to 7 (28%) out of

25 other second-graders who favourably judged a well-intentioned peer. This difference is insignificant using the appropriate chi-square test, $\chi^2(1) = 1.36$.

Judgments of acts involving adult and peer victims

Piaget's notion of adult constraint also implies that younger children, having a unilateral respect for adult authority, should be more likely than older children to believe that to harm an adult is naughtier than to harm a child. But here again this notion is inconsistent with some recent evidence (Siegal, 1979). Children aged 5, 7, 9, and 11 years were asked to compare identical moral acts (lying, stealing, sharing, and helping) committed by a child culprit against either an adult or peer victim. Only the 9-year-olds judged that it is naughtier to harm (and better to help and share with) an adult than a child. The 5- and 7-year-olds displayed no clear preference for either type of act while the 11-year-old group overwhelmingly maintained that both acts are equally naughty (or good) regardless of the age of the victim involved. Thus contrary to the position advocated by Piaget, adult-oriented responses were increasingly evident with age and preceded the mature 'both the same' responses of the 11-year-old group. One interpretation of this result is that the choices of the 11-year-olds may be less the product of the increasing solidarity among friends advocated by Piaget and more the product of a deeper respect for adult authority.

This explanation is further supported by the results of a recent study of children's perceptions of maternal socialisation behaviour (Appel, 1977). Kindergarteners (aged 5 and 6) and third graders (aged 8-10) were given stories in which a highly permissive mother allows a child to misbehave. For example, in one story, a child is permitted to splash others at the beach. The overwhelming majority of third graders evaluated the permissive mother negatively while the kindergarteners overwhelmingly evaluated the mother positively. With age, children appear to develop or 'internalise' a deeper respect for the intentions underlying the adult's child-rearing efforts. An increasing cognitive awareness of such intentions strengthens and legitimises adult authority.

According to Piaget, young children's respect for adult authority is especially manifested in their moral judgments. These disregard the transgressor's intentions and focus more on the consequences of the deed which are 'objectively' defined by adult rules and punishments. Children with age come to give more weight to intentions, a process which according to Piaget is associated with a freedom from adult constraint. But, as we have suggested, an increase in children's understanding of intentionality may produce not a decrease but an increase in the authority of the adult.

Attribution of causes to adult authority

In contrast to the evidence on children's judgments of acts involving adult and peer transgressors and victims, research in the area of children's perceptions of causal attributions has been interpreted to support Piaget's claim that adult authority weakens with age. Karniol and Ross (1976) gave kindergarteners, first graders, second graders, and college students four story pairs. In one member of a pair, for example, a boy played with a toy freely without adult constraint. In the other member, another boy played with an identical toy because his mother told him to do so or promised him a reward. The subjects were then asked, "Who really wants to play with the toy?" The second graders and college students correctly chose the boy who freely played with the toy but the kindergarteners and first graders generally chose the constrained story figure. These results are comparable with those reported in a similar study by Costanzo *et al.* (1974), and are interpreted to be consistent with

Piaget's idea that young children have a unilateral respect for adult authority; an obedient act is better than one which is freely initiated.

However, an opposite explanation is equally plausible. Five-year-olds might judge that the constrained figure is the one who really wants to play with an approved toy only because adult approval implicitly makes a good act better. This does not mean that they respect the adult to the extent that the goodness and badness of acts are dependent on the rewards and punishments dispensed by adults. In fact, Jensen and Hughston (1973) found that 4- and 5-year-olds judge bad acts to be bad independently of whether these acts are rewarded or punished. It might even be the older children who display the greater adult respect here for they are the ones who explicitly recognise that the adult has the power to regulate children's play regardless of their wants and likes.

Before turning to research in other related areas, we believe that a word is in order concerning the methodological problem of ensuring children's story recall and comprehension. For example, in the Karniol and Ross study the stories were audiotaped and presented without visual displays as a memory aid. It is not surprising that, in order for the younger children to remember the second story in a pair, it had to be repeated up to three times. By then they may have forgotten the first story. In addition, the younger children might have chosen the 'constrained' figure because his name was more salient than that of the 'free' figure. Within a story pair, the name of the constrained figure was mentioned six times but the name of the free figure was mentioned only twice.

Having children repeat the stories, even with the benefit of illustrations, is probably not enough of a safeguard to ensure their comprehension. Several studies have indicated that a failure to recall story information can affect children's responses to moral judgment problems (Feldman *et al.*, 1976; Parsons *et al.*, 1976; Austin *et al.*, 1977). For this reason, comprehension probes were used in the Suls and Kalle (1978) study mentioned above to assess children's understanding of the stories. Over 20 per cent of kindergarten and first-grade subjects had to be excluded because of recall errors. Indeed, by going to special lengths to guarantee recall and comprehension, it has been shown that young children can correctly identify a figure who freely chooses a toy as the one who really wants to play in contrast to a constrained figure (Berndt, 1977; Shultz and Butkowsky, 1977).

Effects of adult and peer models on moral judgment development

Thus the evidence both on children's evaluations of adult and peer transgressors and victims and on children's perceptions of causal attributions does not support Piaget's adult constraint notion. Much of this work is inconclusive or suggests the possibility that adult influence on moral judgment increases with age. The same conclusion can be reached through an examination of studies on the comparative effects of adult and peer models on children's moral judgments.

Piaget appears to make no clear prediction on children's susceptibility to external influence over age. While young children are rigid in their advocacy of adult-imposed rules, older children can be meticulous in their adherence to group norms (Allen and Newton, 1972). What Piaget's theory does imply is that older children should be less susceptible to the influence of adult models than are younger children, and that adult models should be less effective than peer models in training children to consider a story character's intentions in judging his moral behaviour.

Yet contrary to Piaget, recent studies have demonstrated that the adult is a generally more effective model than the peer in the case of moral judgment development. Using a design similar to that originally employed by Bandura and McDonald (1963), Dorr and Fey (1974) gave children aged 5 to 11 a pretest consisting of story

pairs which contrasted the behaviour of a child who accidentally caused heavy damage with the light damage caused by a well-intentioned child. On this basis, the subjects were divided into two groups: those who correctly considered the story character's intentions, and those who immaturely disregarded his intentions by focusing on the objective consequences of his behaviour. The groups were then given additional story pairs and exposed to the judgments of either adult or peer models; the models gave objective judgments to the subjective (or intentional) group and intentional judgments to the objective group. In both conditions, the adult model was more influential than the peer model although the latter was more influential than no model.

Similar results were obtained by Brody and Henderson (1977). First graders were simultaneously exposed to two models, a peer and an adult who in some conditions gave conflicting judgments. An adult who gave mature intentional judgments accompanied by rationales was effective in eliciting intentional moral judgments when he was paired with a peer who gave immature objective judgments. An intentional peer model paired with an objective adult was comparatively ineffective.

That adult influence is more powerful than peer influence in stimulating children's moral judgment development is also shown in a study by Lickona (1973, 1976). The judgments of first and second graders were trained under a number of conditions. These included a peer interaction condition involving a face-to-face debate between an objective subject and an already subjective peer and a didactic training condition in which an adult told a subject directly why an intentional judgment is correct. Piaget's theory would imply that the peer interaction condition should be more successful than didactic training. Yet the opposite occurred: didactic training was the more effective of the two. As Lickona points out, this finding is difficult to interpret from a Piagetian standpoint.

Relationship to Kohlberg's theory

At first glance, a possible growth in the influence of adult authority would appear to be entirely consistent with Kohlberg's theory. According to Kohlberg (1976, p. 33), moral development proceeds from a pre-conventional level at which rules and social expectations are something external to the self; to a conventional level at which the self is identified with or has internalised others' rules and expectations, especially those of authorities; and then to a post-conventional level at which the self has been differentiated from others' rules and expectations and values are defined in terms of self-chosen principles.

One of the most useful aspects of Kohlberg's theory is that it attempts to provide a comprehensive account of moral development throughout adolescence and adulthood. His ideas are both highly provocative and controversial (Kurtines and Greif, 1974; Broughton, 1978; Damon, 1978; Philips and Nicolayev, 1979; Rubin and Trotter, 1977; Siegal, in press). But the point which is particularly germane here is that there are differences between the Kohlberg and Piaget theories which pertain to the moral judgment of young children (Kohlberg, 1963, pp. 321-322). For example, contrary to Piaget, Kohlberg claims that 4- and 5-year-olds do not judge acts out of a unilateral respect for adult rules. Instead their judgments can be characterised as pre-conventional and lacking respect for authority. Whether a specific act is externally rewarded or punished is critically important; whether it contravenes an adult rule is insignificant. Thus children should regard a boy who is punished for helping as a bad boy; similarly, a boy who is rewarded for lying should be regarded as a good boy.

Recently this issue has been reopened. In the Jensen and Hughston (1973) study cited earlier, children aged 4 and 5 were given stories in which the central

character performed socially approved, neutral, or negative acts. In half of the stories for each of the three types, the character was punished; in the other half he was rewarded. The findings clearly indicated that the reward outcome had little effect on responses to negative content stories. That is, a naughty deed is evaluated as naughty regardless of whether it is rewarded or punished.

These findings have since been replicated by Gilmore and Ozols (1977). They support Piaget in that young children appear to respect the sacredness of adult rules. But they contradict Kohlberg's claim that 4- and 5-year-olds evaluate an act as bad only if it is punished. It is also noteworthy that nursery school children's judgments about prosocial behaviour also appear to be generally stereotyped and rule-governed rather than oriented toward punishment and obedience (Eisenberg-Berg, 1979; Eisenberg-Berg and Hand, 1979). Perhaps children aged 2 and 3 evaluate acts according to the way described by Kohlberg. But it would be very difficult to test this possibility in such young children.

What is probably needed here is an amalgam of Piaget's and Kohlberg's theories to account for the possibility that adult authority exerts an increasing influence on moral judgment development. Consistent with Piaget, the evidence suggests that the young child has a knowledge and respect for the letter of adult rules. But with age, he begins to appreciate and respond to the spirit of these rules, and to the intentions underlying their implementation. At this point, consistent with Kohlberg, the child develops a conventional morality which involves a recognition that adult rules are intentionally aimed to prevent naughty or unjust behaviour.

CONCLUSIONS

Though the evidence is indirect in the sense that it has been derived almost wholly from experimental as opposed to observational studies, findings from divergent research areas converge in pointing to the same conclusions. Our review has indicated little support for Piaget's claim that children's respect for adult authority is liberated with increasing peer group interaction. Children of all ages judge adult and peer transgressors as equally naughty. With age, children in fact appear to develop a deeper respect for the adult victim and the intentions underlying the adult's efforts at child care. In the realm of social perception, children also may come to recognise explicitly that the adult has the power and authority to control children's behaviour regardless of their wants and likes. Moreover the adult is more effective than the peer as a model for stimulating moral judgment development. Children's susceptibility to adult influence may well decline for perceptual judgments of unambiguous stimuli and on matters of opinion (Utech and Hoving, 1969; Allen and Newton, 1972). But as children grow older they become more adept at recognising that adult rules are deliberately designed to prevent naughty behaviour. The child can identify more strongly with these rules because he recognises their rational quality.

This is a possibility which Piaget considers but then dismisses as characteristic of primitive non-Western societies: "In our societies the child of 13 escapes from the family circle and comes in contact with an ever-increasing number of social circles which widen his mental outlook. Whereas in so-called primitive communities adolescence is the age of initiation, therefore, of the strongest moral constraint, and the individual as he grows older, becomes more and more dependent. But keeping in mind only our societies of children we see that co-operation constitutes the most deep-lying social phenomenon, and that which has the surest psychological foundations. As soon as the individual escapes from the dominion of age, he tends towards co-operation as the formal form of social equilibrium" (Piaget, 1932/1977, p. 99).

There is little doubt that the increasingly complex moral dilemmas of modern

society render Piaget's claim problematic and questionable. An age of initiation exists not only in economically underdeveloped societies but also in the industrialised West. In order to solve these moral problems, a learner, whether child or adolescent, has to be 'initiated' into skills of reasoning (Peters, 1967). Education is a process of initiation and is facilitated by an increasing respect on the part of the child learner for the adult teacher. A lack of respect can impair this process, and evidence suggests that children's readiness to misbehave is associated with disrespectful attitudes towards adults and is comparatively unassociated with attitudes towards peers (Bixenstine *et al.*, 1977).

To be sure, Piaget (1932/1977, p. 190) does allow that children's moral development is "naturally capable of being reinforced by the precepts and the practical example of the adult" though it is "largely independent of these influences". But he claims that traditional educational practice which directly attempts to inculcate adult values is dangerous in that it discourages co-operation and encourages an unreflecting conformity to adult rules. A "new", more original method of education should centre on developing co-operation within the peer group (Piaget, 1971, pp. 178-180).

Yet if adult authority does not crumble beneath the onslaught of peer influence it is especially incumbent upon the adult himself to develop children's sense of right and wrong. To this end, peer group debates in the classroom have often been used to stimulate children's moral judgment development (Blatt and Kohlberg, 1975; Jensen and Murray, 1978; Power and Reimer, 1978). But perhaps less emphasis should be placed on peer group centred programmes and more stress put on the effectiveness of direct adult intervention. This of course is not to deny that peer group influence can be profitably channelled toward increasing, for example, children's communication skills (Gottman *et al.*, 1975) or their disposition to act prosocially (Staub, 1979). But at least in the area of moral judgment development, the adult himself may actively and systematically encourage children to imagine the interests and intentions of others through modelling, reinforcement, and reasoning. Given that elementary school teachers can have an important effect on the child's subsequent adult status (Pederson and Faucher, 1978), this directed effort may be worthwhile and lasting.

ACKNOWLEDGMENTS.—This paper was prepared with the support of the Humanities and Social Sciences Research Grant Committee of the University of British Columbia. Thanks are due to Dr. Michael Chandler and Dr. Dale Miller for critical comments on an earlier version of the manuscripts.

REFERENCES

- ALLEN, V. L., and NEWTON, D. (1972). Development of conformity and independence. *J. pers. soc. Psychol.*, 22, 18-30.
- APPEL, Y. H. (1977). Developmental differences in children's perception of maternal socialization behavior. *Child Dev.*, 48, 1689-1693.
- AUSTIN, V. D., RUBLE, D. N., and TRABASSO, T. (1977). Recall and order effects as factors in children's moral judgments. *Child Dev.*, 48, 470-474.
- BANDURA, A., and McDONALD, F. J. (1963). Influence of social reinforcement and the behavior of models in shaping children's moral judgments. *J. abn. soc. Psychol.*, 67, 241-281.
- BERNDT, T. J. (1977). The effect of reciprocity norms on moral judgment and causal attribution. *Child Dev.*, 48, 1322-1330.
- BIXENSTINE, V. E., DECORTE, M. S., and BIXENSTINE, B. A. (1977). Conformity to peer-sponsored misbehavior at four age levels. *Dev. Psychol.*, 12, 226-244.
- BLATT, M. M., and KOHLBERG, L. (1975). The effects of classroom moral discussion upon children's level of moral judgement. *J. moral Educ.*, 2, 129-161.
- BRODY, G. H., and HENDERSON, R. W. (1977). Effects of multiple model variations and rationale provision on the moral judgments and explanations of young children. *Child Dev.*, 48, 1117-1120.
- BROUGHTON, J. M. (1978). The cognitive-developmental approach to morality: a reply to Kurtines and Greif. *J. moral Educ.*, 8, 81-96.

- COSTANZO, P. R., GRUMET, J. F., and BREHM, S. S. (1974). The effects of choice and source of constraint on children's attribution of preference. *J. exp. soc. Psychol.*, 10, 352-364.
- DAMON, W. (1978). *Moral Development: New Directions for Child Development, Number 1*, San Francisco: Jossey-Bass.
- DORR, D., and FEY, S. (1974). Relative power of symbolic adult and peer models in the modification of children's moral choice behavior. *J. pers. soc. Psychol.*, 29, 335-341.
- EISENBERG-BERG, N. (1979). Development of children's prosocial moral judgment. *Dev. Psychol.*, 15, 128-137.
- EISENBERG-BERG, N., and HAND, M. (1979). The relationship of preschoolers' reasoning about prosocial moral conflicts to prosocial behavior. *Child Dev.*, 50, 356-363.
- FELDMAN, N. S., KLOSSON, E. C., PARSONS, J. E., RHOLES, W. S., and RUBLE, D. N. (1976). Order of information presentation and children's moral judgements. *Child Dev.*, 47, 556-559.
- GILMORE, J. B., and OZOLS, E. (1977). Respect rules hedonism. Paper presented at the biennial meeting of the Society for Research in Child Development, New Orleans.
- GOTTMAN, J., GONSO, J., and RASMUSSEN, B. (1975). Social interaction, social competence, and friendship in children. *Child Dev.*, 46, 709-718.
- JENSEN, L. C., and HUGHSTON, K. (1973). The relationship between type of sanction, story content, and children's judgments which are independent of sanction. *J. gen. Psychol.*, 122, 49-54.
- JENSEN, L. C., and MURRAY, M. (1978). Facilitating development of four moral concepts among kindergarten and first grade children. *J. educ. Psychol.*, 70, 936-944.
- KARNIOL, R. (1978). Children's use of intention cues in evaluating behavior. *Psych. Bull.*, 85, 76-86.
- KARNIOL, R., and ROSS, M. (1976). The development of causal attributions in social perception. *J. pers. soc. Psychol.*, 34, 455-464.
- KEASEY, C. B. (1977). Young children's attribution of intentionality to themselves and others. *Child Dev.*, 48, 261-264.
- KOHLBERG, L. (1963). The development of children's orientations toward a moral order. I. Sequence in the development of moral thought. *Human Dev.*, 6, 11-33.
- KOHLBERG, L. (1976). Moral stages and moralization: The cognitive developmental approach. In LICKONA, T. (Ed.), *Moral Development and Behavior*. New York: Holt, Rinehart and Winston.
- KOHLBERG, L. (1978). Revisions in the theory and practice of moral development. In DAMON, W. (Ed.), *Moral Development: New Directions for Child Development, Number 1*. San Francisco: Jossey-Bass.
- KURTINES, W., and GREIF, E. B. (1974). The development of moral thought: review and evaluation of Kohlberg's approach. *Psych. Bull.*, 81, 453-478.
- LICKONA, T. (1973). An experimental test of Piaget's theory of moral development. Paper presented at the biennial meeting of the Society for Research in Child Development, Philadelphia.
- LICKONA, T. (1976). Research on Piaget's theory of moral development. In LICKONA, T. (Ed.), *Moral Development and Behavior*. New York: Holt, Rinehart and Winston.
- PARSONS, J. E., RUBLE, D. N., KLOSSON, E. C., FELDMAN, N. S., and RHOLES, W. S. (1976). Order effects on children's moral and achievement judgements. *Dev. Psych.*, 12, 357-358.
- PEDERSON, E., and FAUCHER, T. A. (1978). A new perspective on the effects of first-grade teachers on children's subsequent adult status. *Harvard J. Educ.*, 48, 1-31.
- PETERS, R. S. (1967). What is an educational process? In PETERS, R. S. (Ed.), *The Concept of Education*. London: Routledge and Kegan Paul.
- PETERSON, C., PETERSON, J., and FINLEY, N. (1974). Conflict and moral judgment. *Dev. Psych.*, 10, 65-69.
- PHILIPS, D. C., and NICOLAYEV, J. (1979). On assessing Kohlberg's stage theory of moral development. In COCHRANE, D. B., HAMM, C. M., and KAZEPIDES, A. C. (Eds.), *The Domain of Moral Education*. New York: The Paulist Press.
- PIAGET, J. (1977). *The Moral Judgement of the Child*. Harmondsworth, England: Penguin (originally published London: Routledge and Kegan Paul, 1932).
- PIAGET, J. (1971). *Science of Education and the Psychology of the Child*. New York: Penguin.
- POWER, C., and REIMER, J. (1978). Moral atmosphere: an educational bridge between moral judgment and action. In DAMON, W. (Ed.), *Moral Development: New Directions for Child Development, Number 2*. San Francisco: Jossey-Bass.
- RUBIN, K. H., and TROTTER, K. T. (1977). Kohlberg's moral judgment scale: some methodological considerations. *Dev. Psychol.*, 13, 535-536.
- RYBASH, J. M., SEWALL, M. B., ROODIN, P. A., and SULLIVAN, L. (1975). Effects of age of transgressor, damage, and type of presentation on kindergarten children's moral judgments. *Dev. Psychol.*, 11, 874.
- SCHULTZ, T. R., and BUTKOWSKY, I. (1977). Young children's use of the scheme for multiple sufficient causes in the attribution of real and hypothetical behavior. *Child Dev.*, 48, 464-469.
- SIEGAL, M. (1975). Spontaneous development of moral concepts. *Human Dev.*, 18, 370-383.
- SIEGAL, M. (1979). Socialization and the development of adult respect. *Br. J. Psychol.*, 70, 83-86.
- SIEGAL, M. (in press). Kohlberg vs. Piaget: to what extent has one theory eclipsed the other? *Merrill-Palmer Q.*

- STAUB, E. (1979). *Positive Social Behavior and Morality. Volume 2: Socialization and Development*. New York: Academic Press.
- SULS, J., and KALLE, R. J. (1978). Intention, damage, and age of transgressor as determinants of children's moral judgments. *Child. Dev.*, 49, 1270-1273.
- THOMSON, J. A. K. (1953). *The Ethics of Aristotle*. Harmondsworth, England: Penguin.
- UTECH, D. A., and HOVING, K. L. (1969). Parents and peers as competing influences on the decisions of children of differing ages. *J. soc. Psychol.*, 78, 267-274.

(Manuscript received 17th August, 1979)

SELF-REGARD OF PRIMARY SCHOOL CHILDREN: SOME AUSTRALIAN DATA

By D. A. ELLERMAN

(*Darling Downs Institute of Advanced Education, Toowoomba, Queensland*)

SUMMARY. Samples of children ($N = 1,267$) attending primary schools in an Australian rural setting completed the *Where Are You Game* for the assessment of self-regard, in order to investigate the test's psychometric properties and the correlates of test scores. Results showed the test to yield meaningful scores over the full age-range of primary school, and to have acceptable estimated reliability. Factor analysis indicated the test to be uni-factorial in structure. Correlations with observer ratings of pupil characteristics showed considerable convergent validity. Analysis of variance indicated that younger children had more favourable self-regard than older children, those with good school achievement saw themselves more favourably than low achievers, the boys more favourably than girls. It is concluded that the test shows promise as a research instrument.

INTRODUCTION

SINCE the 1940s there has been a vast amount of research effort in the area of self-concept theory. Wylie (1974) estimated over 2,000 publications appeared between 1961 and 1971 alone. There is great diversity in the ways in which the term *self* has been used, varying roughly between usages implying self-as-object of a person's own knowledge and evaluation, and self-as-process or agent. The present study, in common with most others, is based on the conception of self-as-object, but implying also an active role in influencing behaviour in some aspects.

An important part of the field of self-concept theory is that of the development of self-regard in both adults and children, and the factors which influence its formation. There have been few large-scale or intensive studies which attempt to chart systematically the development of self-esteem with increasing age or in relation to the primary socialisation influences of family, peer group, school experiences, sex-role differences and so on. Most studies have been 'one-shot', and unrelated to a systematic programme of research.

Amongst the many frequently studied correlates of self-regard have been academic achievement, family socio-economic status, age and sex differences, peer popularity and many others. Yet even on such basic variables, results are conflicting. For example, the review of Purkey (1970) and studies by Lamy (1965) Coopersmith (1967), Caplan (1969), and Cummings (1971) found a significant, though generally weak, relationship between self-concept and academic achievement level, while others—Badwal (1969), La Belle (1970), Beebe (1972), Chang (1976)—found no significant differences.

A similar pattern of conflicting results occurs for sex differences. Carpenter and Busse (1969) found girls to have more negative self-concepts. Wylie (1961) came to a similar conclusion at least for tests involving real/ideal self measures. Piers and Harris (1964) reported no significant sex differences, nor did Sears (1970) on five different measures of self-concept. Coopersmith (1959), on the other hand, reported a slight though non-significant trend for females to score higher than males.

Age or grade level differences have been charted systematically in fewer studies. Morse (1964) found lower self-esteem in grade 11 than in grade 3. Piers and Harris (1964) found grade 6 children to have lower scores than those in both grade 3 and grade 10. Chang (1976) on the other hand reported no significant differences between 4th, 5th and 6th grade children in her sample.

It is probable that a large part of the confusion over these basic variables can be explained partly by the uncertain validity and differing item content of the measures of self-esteem used. The most frequently used procedures, such as those of Lipsitt (1958), Coopersmith (1967), Piers (1972) and Bills (undated), all have areas in which validity is not yet fully researched (Wylie, 1974). All require considerable reading fluency from subjects. The only widely used non-verbal test of Long *et al.* (1970) has yet to substantiate the claimed special advantages of the method, and to provide data on the special validity problems which are associated with the test (Wylie, 1974, p. 198).

One test which has been used in a few studies, but which appears promising from several of its characteristics, is the *Where Are You Game* (WAYG) of Engel and Raine (1963). It is adaptable for use with pre-literate children. It is simple enough for an untrained tester to administer. It is brief. For these reasons the WAYG seems to hold promise as a potential research instrument. Hence it was selected for further investigation.

The purposes of the present study were twofold: (a) to investigate further the psychometric properties of the WAYG; and (b) to examine the differences in levels of overall self-regard associated with the basic variables of sex, grade level and academic achievement.

METHOD

Subjects

Three distinct samples of primary school children in grades 1 to 7 were selected and are described in detail below ($N = 1,267$). All subjects were attending primary schools in the south-east region of Queensland, Australia, mostly in the area of Toowoomba, an inland rural city with a population of approximately 60,000. Subjects were selected according to differing instructions given to the interviewers of each sample, in consultation with classroom teachers.

Interviewers were all students enrolled in a Diploma of Teaching course. The interviews and a subsequent report were required parts of course assessment.

Sample 1 ($N = 553$) was selected in October and November of 1975, while interviewers were assigned to the classroom for a period of practical experience. Each interviewer was required to select four pupils: one boy whose academic achievement was above average for the grade and one who was below average, and two similarly contrasting girls. Thus the sampling strategy was designed to achieve a balanced representation of academic achievement categories.

Sample 2 ($N = 567$) was selected for interview in October, 1976. In this sample interviewers were required to select two boys and two girls at random from the classroom to which they were assigned for a period of practical experience.

Sample 3 ($N = 147$) was selected in June and July, 1977. Interviewers were required to select one child only for more detailed study, and it was specified that it should be a child they knew quite well, either from their in-school experience period the previous term or, if this was not possible, from any other social contact (friends, neighbours, etc.) available to them. Approximately 42 per cent (61 subjects) were attending schools outside the local Toowoomba region. For this sample only, interviewers also completed the Health Resources Inventory of Gesten (1976). Sample characteristics are described in Table 1.

Materials

Self-regard was assessed by means of the *Where Are You Game* (WAYG) created by Engel and Raine (1963), with format similar to that described by Carpenter and

TABLE 1
SEX COMPOSITION OF SAMPLES USED

	Boys	Girls	Total
Sample 1	287	266	553
Sample 2	279	286	567
Sample 3	67	79	147
Total	633	631	1267

NOTE: Sex was not recorded for two subjects in Sample 2 and one in Sample 3.

Busse (1969) and Busse *et al.* (1974). In the present version children are asked to rate themselves on each of seven dimensions. These are seeing oneself as:

- (a) happy vs. sad
- (b) smart vs. not smart
- (c) well liked vs. not well liked
- (d) athletic vs. non-athletic
- (e) attractive vs. unattractive
- (f) strong vs. weak
- (g) obedient vs. disobedient

These scales are the same as those used for the earliest study with minor variation in wording, except for item (d) which replaces the original *brave vs. afraid*. The seven scales are printed vertically on a single page, with both of the poles of the dimension labelled at the extremes with the concept labels, together with a stick figure of indeterminate sex used in the explanation of the responses required from the child. Labelling is counterbalanced in order (*happy vs. sad*, followed by *not smart vs. smart*, etc.). Scoring is on a 5-point scale for each dimension with each point corresponding to one of five boxes between the extreme poles, a score of 1 indicating the box closest to the less positive pole (e.g., *sad*), and 5 indicating a score closest to the positive pole (e.g., *happy*). Total scores are the sum of ratings for each dimension, giving a possible score range from the lowest of 7 to the highest of 35.

Personal and social competence. The Health Resources Inventory (HRI) of Gesten (1976) is a relatively new scale for use by teachers to make global assessment of primary school children's competency-related behaviour and is designed to measure the positive aspects of mental health in children. Gesten reports a replicated structure of five factors for the HRI. (I = Good Student; II = 'Gutsy'; III = Peer Sociability; IV = Rules; V = Frustration Tolerance.) Gesten's original inventory was made up of 54 items. The present study, however, makes use of a shortened version, including only the 44 items which had high factor score coefficients (>0.10) on one of the five factors. For each item, the child is rated according to how well the item describes that child, using a 5-point scale from 1 = not at all, to 5 = very well.

Procedure

Interviewers for all three samples followed the same basic procedure. After each subject was selected according to the varying criteria and methods outlined above, the interviewer took the child to a room where they would be free of distraction and disturbance to complete the WAYG Self-Regard Scale. After a short period of informal talk aimed at building rapport, interviewers said they were going to play a game. A full explanation of the way the children should make ratings of themselves on the seven dimensions was given, making use of the two stick figures on the poles

of each dimension to aid in explanation. The schedule was standardised for all interviews.* Interviewers reported that for younger children (grades 1-2) some rephrasing or further explanation of how responses were to be made was occasionally necessary, but most children had no difficulty in making apparently meaningful responses.

In addition to the self-regard data, each interviewer recorded data on school attended, sex, age, grade level, and an estimate of the child's overall academic performance compared to all children in the same grade in that school. Interviewers consulted classroom teachers and school records to make this judgment and simply classified the child into one of four 'quartile' groups—top 25 per cent (good achievement); between 50 per cent and 75 per cent (above average); between 50 per cent and 25 per cent (below average); and lowest 25 per cent (poor achievement).

Finally, the child was thanked for his or her co-operation, and the interviewer then departed. Total time for each interview varied between 5-20 minutes. For sample 3 only, the interviewers subsequently completed the HRI within a short time of the interview.

RESULTS

Initial analysis revealed no significant differences between samples 1 and 2 (the two large-scale samples) on item or total self-regard scores. Hence these have been combined for most analyses, and are referred to as the main sample ($N = 1,120$).

Distribution of scores on the WAYG

Table 2 presents the mean total scores and standard deviations for the main sample classified by sex, grade and achievement. These data may be taken as tentative norms for research purposes in comparable settings. Differences in total scores were tested by three separate one-way analyses of variance.

Analysis by Grade level indicated a highly significant difference ($F = 50.8$, $df = 6/1113$, $P < 0.001$). Scheffé's procedure (Glass and Stanley, 1970, pp. 388-393) was used to determine where the significant differences between the seven means lay. It was found that grades four to seven formed a homogeneous subset, within which means did not differ from each other, as did grades one and two ($\alpha = 0.05$). Remaining comparisons were all significant. Thus children in the first two grades had the most positive self-regard, and children in grade four onwards the least positive, with those in grade three occupying a middle position.

One-way analysis by Achievement also revealed highly significant differences between groups ($F = 11.6$, $df = 3/1101$, $P < 0.001$). Scheffé's test indicated that there were no significant differences between the poor, below average or above average groups; the good achievers (top 'quartile' group) had significantly more positive total self-regard scores than the others. The final one-way analysis by Sex indicated that boys had significantly higher total scores than girls ($F = 4.0$, $df = 1/1116$, $P < 0.05$).

Possible interactions between these three main effects were tested by a three-way ($3 \times 4 \times 2$) analysis of variance using Dixon's (1975) programme. Grades were collapsed into three levels (1, 2 vs. 3, 4, 5, vs. 6, 7). Only one interaction was significant: that for Grade $\times$ Achievement ($F = 3.0$, $df = 6/1079$, $P < 0.006$). From the table of means it can be seen that this reflects the trend for a growing disparity between the scores for high and low achievers at increasing grade level, with low achievers having increasingly negative self-regard.

* Copies available from the author.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF WAYG TOTAL SCORES FOR SUBJECTS CLASSIFIED BY SEX,
GRADE AND ACHIEVEMENT LEVEL

		Grade level						
Group		1	2	3	4	5	6	7
<i>Boys</i>								
Good	M	30.2	29.4	28.4	26.7	26.3	27.1	26.1
	SD	3.2	3.3	4.8	4.6	3.9	5.4	4.0
	N	23	26	17	22	32	20	37
Above ave.	M	29.8	29.0	26.4	25.9	23.8	25.3	23.5
	SD	4.3	5.1	2.2	2.8	4.2	3.2	3.9
	N	23	18	16	14	21	29	28
Below ave.	M	29.2	27.5	28.0	23.3	24.9	23.5	23.9
	SD	4.4	5.9	4.6	4.1	5.1	2.2	4.2
	N	20	10	9	12	14	19	35
Poor	M	32.4	29.9	27.7	25.3	24.6	23.3	22.7
	SD	3.4	4.3	4.7	3.9	4.7	3.8	4.0
	N	12	19	15	15	16	21	18
Total	M	30.2	29.2	27.6	25.5	25.1	24.9	24.3
	SD	4.0	4.4	4.1	4.1	4.4	4.0	4.2
	N	78	73	57	63	83	89	118
<i>Girls</i>								
Good	M	29.3	29.8	27.8	25.8	25.3	24.7	25.9
	SD	5.2	3.6	4.7	2.9	4.6	4.7	3.6
	N	33	25	25	21	28	22	38
Above ave.	M	29.5	27.6	27.8	24.9	23.5	23.9	24.2
	SD	5.5	4.4	4.6	4.7	3.2	4.0	4.1
	N	20	20	13	15	18	22	35
Below ave.	M	28.9	30.7	26.0	23.9	24.3	24.6	21.5
	SD	4.3	5.1	2.9	4.3	4.0	3.0	3.1
	N	9	12	10	12	22	17	25
Poor	M	30.0	29.9	25.9	20.7	20.6	23.2	22.7
	SD	6.4	3.8	5.5	6.7	4.2	4.4	6.2
	N	14	12	14	9	11	20	20
Total	M	29.4	29.3	27.1	24.3	23.9	24.1	23.9
	SD	5.4	4.2	4.6	4.6	4.3	4.1	4.5
	N	76	69	62	57	79	81	118

Note: N=1103. Sex, grade or achievement not recorded for 93 subjects.

Properties of the WAYG

Internal consistency. Internal consistency was assessed for each sample separately and for the main sample, using the computer programme provided by Specht (1976). Cronbach *alphas* ranged between 0.72 for sample 1 to 0.77 for sample 2. For the main sample *alpha* was 0.75. Though not high, these estimates provide evidence of a reasonable level of internal consistency for the test.

Inter-item correlations: Pearson *rs* for the correlations between items are shown in Table 3. All *rs* are moderately positive, ranging from 0.19 to 0.37.

Factor analysis of this matrix (following Veldman, 1967) produced a factor solution with only one eigenvalue greater than 1.0, which accounted for 40.1 per cent of the total variance. All seven items had factor loadings between 0.56 and 0.70 on this factor. Analysis for each of the three samples separately produced the same unifactorial solution with variance accounted for ranging from 38 per cent to 44 per cent. Corrected item-total correlations are also presented in Table 3 and are moderately positive in the range from 0.40 to 0.53.

TABLE 3
INTER-ITEM AND ITEM-TOTAL CORRELATIONS (N = 1120)

Item	Happy	Smart	Well-liked	Athletic	Attractive	Strong	Obedient
Smart	0.29						
Well-liked	0.34	0.28					
Athletic	0.24	0.19	0.31				
Attractive	0.32	0.34	0.37	0.29			
Strong	0.29	0.27	0.32	0.32	0.37		
Obedient	0.31	0.34	0.27	0.21	0.35	0.26	
Corrected Item-Total	0.46	0.44	0.49	0.40	0.53	0.48	0.44

Self-ratings and interviewer ratings

For subjects in sample 3, interviewers completed ratings for 44 items of the Health Resource Inventory after administering the WAYG to the child. Resulting HRI items were then subjected to factor analysis. In the present shortened version, a five-factor solution was substantially the same as Gesten's (1976) original report. The first varimax rotated factor was almost identical with the original *Rules* factor. Factor II contained all but one of the items with high loadings on the original *Frustration Tolerance* factor. This factor is of particular interest, since it has most of the items with self-referent content. Indeed, it could well have been renamed Self-Esteem. The item with the highest loading was *feels good about himself or herself*. Other high loading items (in order of loadings) included *generally relaxed, accepts things not going own way* and *approaches new experiences confidently*.

The third factor is almost identical with Gesten's *Good Student* factor. Factor IV contained mostly items from the original Peer Sociability factor plus four items from the 'Gutsy' factor, and can be still adequately described as *Peer Sociability*. Factor V, however, did not correspond closely to any of those in Gesten's original structure, but seemed to be concerned with being self-reliant or *Independent*.

The correlations of individual factor scores with the items and total scores on the WAYG are presented in Table 4.

Total scores show highly significant correlations with both Factor II (*Frustration*

TABLE 4
CORRELATIONS OF WAYG ITEMS AND TOTAL SCORES WITH HRI FACTOR SCORES (N = 147)

WAYG item	HRI factor scores				
	I (Rules)	II (Frustration Tolerance)	III (Good Student)	IV (Peer Sociability)	V (Independent)
Happy	-0.06	0.34*	0.21	0.05	0.05
Smart	-0.08	0.23	0.45**	0.09	-0.07
Well-liked	-0.07	0.37**	0.23	0.18	0.07
Athletic	-0.05	0.21	0.25	0.06	0.09
Attractive	0.00	0.39**	0.18	0.18	0.07
Strong	-0.17	0.22	0.16	-0.04	0.02
Obedient	0.24	0.11	0.24	-0.06	-0.04
Total	-0.04	0.43**	0.40**	0.11	0.04

* $P < 0.05$, ** $P < 0.01$

Tolerance) and Factor III (*Good Student*). Correlations with other factor scores are low.

The patterns of correlations of factor scores with individual WAYG items for Factors II and III, show consistency of magnitude and direction, and range up to 0.45 for *smart* with *Good Student* (Factor III). Correlation of WAYG items with scores for Factors I, IV and V are all low. Predictable, though non-significant relationships are found between WAYG self-rating of *obedient* and the *Rules* factor ($r = 0.24$) and between *well-liked* and *attractive* with the *Peer Sociability* factor (both $r_s = 0.18$).

DISCUSSION

In general it is perhaps reassuring to find that these samples of Australian rural children had quite positive self-evaluations with the overall mean score of around 26 for the main sample being well above 21, the mid-point of the possible score range.

This finding contrasts somewhat with the results of Carpenter and Busse (1969) for their low SES American sample, where the mean scores for their first and fifth grade students were both about five points lower than in the present study. By way of explanation, one can only speculate on differences in culture, socio-economic status, time and setting (urban vs. rural) or other variables which differ between the samples.

Psychometric properties

The analyses of the psychometric properties of the WAYG reported here seem to provide an initial basis for concluding that it is a useful research instrument for use with children over the full range of primary school ages from 5 to 12 years. It has the attraction of being both brief and simple to use even for relatively unskilled testers. The test has an acceptable level of internal consistency of around 0.75. This compares well with the estimated reliability of the original version reported by Engel and Raine (1963) of 0.60. Factor analysis of the present version revealed a simple unifactorial structure, in contrast to the complex five-factor solution presented by Engel and Raine, though this difference may be due to the more conventional and conservative criteria used for factor extraction in the present study.

Analysis of the relationships between both item and total scores on the WAYG with HRI factor scores provides some evidence for the convergent validity of the test. Since the WAYG is a measure of phenomenal or subjective self-regard, and the HRI is a behavioural rating made by interviewers, high correlations cannot be expected. The *Frustration Tolerance* factor is that with the greatest face validity in relation to self-regard. The highly significant correlation with total WAYG scores confirms this. In addition, the highly significant correlation of the three WAYG items most concerned with satisfaction in interpersonal relationships (*attractive, well-liked, happy*) suggests that there may be some possible causal link between the skills necessary to cope with stress, maintain emotional stability and have a favourable self-image contained in this HRI factor, and the subjective enjoyment of interpersonal relationships through admiration by and popularity with others.

The second highly significant correlation of total WAYG scores is with Factor III (*Good Student*). Again this provides evidence for the convergent validity of the test if the assumption is made of some causal link between academic success and increased self-regard, as suggested by Purkey (1970) and others.

Developmental implications

Three significant effects are reported for differences in grade level, sex and academic achievement. Of these, the strongest overall trend is for self-regard to

decline in favourability with increasing age. This accords with the mean difference of 5.7 points noted by Carpenter and Busse (1969) in comparing grade 1 and grade 5 welfare children. This trend may be a function of many influences including increasingly 'accurate' self-perception, growing influence of a shared modesty norm concerning self-disclosure to relative strangers, an increasing ability to make differentiated responses to such self-report tasks, or a multitude of environmental socialising influences impinging on the child. The importance of these or other possible explanations cannot be assessed with the present data. Indeed, with a cross-sectional study such as the present one, it is not possible to be completely certain that individuals' scores do in fact decline without conducting the necessary longitudinal panel study, though the alternative explanations of sampling bias or changes in age-cohort experience seem highly unlikely.

The significant difference in total WAYG scores associated with reported academic achievement is perhaps the most interesting of the findings. Taken together with the significant interaction with grade level, it implies a development sequence by which children who are doing poorly in school increasingly come to hold a more negative view of themselves, through repeated experience of failure. As Purkey (1970) suggests, this is probably a complex two-way relationship, with those who are doing well in school not only internalising a positive view of their abilities but also enjoying more satisfactory relationships with peers, teachers and parents as a result of their success, and thus receiving much more social reinforcement. This in turn increases the child's motivation to approach academic tasks with confidence, and to persist longer at them, since they have learned to expect support and assistance from others when difficulty is found. This view is supported by a supplementary analysis of individual items which indicated that the significant effects for Achievement are found not only with the clearly intellectual/academic item *smart*, but also for those concerning interpersonal relations as well—*well-liked*, *attractive*, *obedient* and *happy*. There were no significant differences for the items *athletic* and *strong*, which might be characterised as concerned with physical skills.

The sex differences reported here are in accord with those of Carpenter and Busse (1969) in finding less positive overall ratings for girls than boys. As noted above, other studies in this area are contradictory. The present data suggest lower self-esteem for girls, but given the sex-role related content of some individual items, it is not possible to conclude whether this is a reflection of sex-role identification, time selection, internalisation of stereotypical derogation of feminine traits, or some other process.

In conclusion, the WAYG appears to be a valid and reliable test for the assessment of children's self-regard for research purposes. It is suitable for a wide age-range of children and can be used by untrained interviewers. Despite this promising progress report, there is much work still to be done. A full multitrait-multimethod validation study is needed. Test-retest reliability needs to be assessed, as does the influence of factors such as social desirability response patterns. Possible expansion of the age range both upwards and downwards would extend the test's usefulness. Further expansion of the content to include items more favourable to the female sex-role, and perhaps other areas, may provide a more comprehensive assessment of global self-regard.

REFERENCES

- BADWAL, B. S. (1969). A study of the relationship between attitude towards school and achievement: sex and grade level. *Dissert. Abstr.*, 30, 6-A, 2366.
BEEBE, J. D. (1972). Self concept and achievement among elementary students in an experimental program. Unpublished Doctoral Dissertation, University of North Dakota.
BILLS, R. E. (no date). *Index of Adjustment and Values. Forms Elementary, Junior High School, and High School: Manual*. University of Alabama: mimeograph.

- BUSSE, T. V., MANSFIELD, R. S., and MESSINGER, L. J. (1974). *Activities in Child and Adolescent Development*. New York: Harper and Row.
- CAPLIN, M. D. (1969). The relationship between self-concept and academic achievement. *J. exp. Educ.*, 37, 13-15.
- CARPENTER, T. R., and BUSSE, T. V. (1969). Development of self-concept in negro and white welfare children. *Child Dev.*, 40, 935-939.
- CHANG, T. S. (1976). Self concepts, academic achievement and teacher ratings. *Psychol. in Schools*, 13, 111-113.
- COOPERSMITH, S. (1959). A method for determining types of self-esteem. *J. abnorm. soc. Psychol.*, 59, 87-94.
- COOPERSMITH, S. (1967). *The Antecedents of Self-Esteem*. San Francisco: Freeman.
- CUMMINGS, R. N. (1971). A study of the relationship between self-concepts and reading achievement at third-grade level. *Dissert. Abstr.*, 31, 10-A, 5195.
- DIXON, W. J. (1975). *BMDP Biomedical Computer Programs*. Los Angeles: University of California Press.
- ENGEL, M., and RAINE, W. J. (1963). A method for the measurement of the self-concept of children in the third grade. *J. gen. Psychol.*, 102, 125-137.
- GESTEN, E. L. (1976). A Health Resources Inventory: the development of a measure of the personal and social competence of primary-grade children. *J. consult. clin. Psychol.*, 44 (5), 775-786.
- GLASS, G. V., and STANLEY, J. C. (1970). *Statistical Methods in Education and Psychology*. Englewood Cliffs: Prentice Hall.
- LA BELLE, T. J. (1970). Attitudes and academic achievement among male and female Anglo- and Spanish-American fifth-grade students. *Dissert. Abstr.*, 31, 4, 1624-A.
- LAMY, M. W. (1965). Relationship of self-perceptions of early primary children to achievement in reading. In GORDON, I. J. (Ed.), *Human Development: Readings in Research*. Chicago: Scott Foresman.
- LIPSITT, L. P. (1958). A self-concept scale for children and its relationship to the children's form of the Manifest Anxiety Scale. *Child Dev.*, 29, 463-472.
- LONG, B. H., HENDERSON, E. H., and ZILLER, R. C. (1970). *Manual for the Self-Social Symbols Tasks and the Children's Self-Social Constructs Tests*. Mimeographed. Baltimore: Goucher College.
- MORSE, C. W. (1964). Self-concept in the school setting. *Childhood Educ.*, 41, 195-198.
- PIERS, E. V. (1972). Parent prediction of children's self-concepts. *J. consult. clin. Psychol.*, 38, 428-433.
- PIERS, E. V., and HARRIS, D. B. (1964). Age and other correlates of self-concept in children. *J. educ. Psychol.*, 55, 91-95.
- PURKEY, W. W. (1970). *Self-Concept and School Achievement*. Englewood Cliffs, N.J.: Prentice Hall.
- SEARS, R. R. (1970). Relation of early socialization experiences to self-concepts and gender role in middle childhood. *Child Dev.*, 41, 267-289.
- SPECHT, D. A. (1976). *SPSS: Statistical Package for the Social Sciences: Version 6 Users Guide to Subprogram Reliability and Repeated Measures Analysis of Variance*. Dept. of Sociology, Iowa State University.
- VELDMAN, D. J. (1967). *Fortran Programming for the Behavioral Sciences*. New York: Holt, Rinehart and Winston.
- WYLIE, R. C. (1961). *The Self-Concept: A Critical Review of Pertinent Research Literature*. Lincoln: University of Nebraska Press.
- WYLIE, R. C. (1974). *The Self-Concept: A Review of Methodological Considerations and Measuring Instruments*. (Rev. edit.) Vol. 1. Lincoln: University of Nebraska Press.

(Manuscript received 31st October, 1979)

THE KEYWORD METHOD AND CHILDREN'S VOCABULARY LEARNING IN THE CLASSROOM

By R. MERRY

(*School of Education, University of Leicester*)

SUMMARY. In an attempt to relate recent experimental findings of the effectiveness of the keyword technique in vocabulary learning more closely to actual classroom practice, an exploratory questionnaire suggested that many teachers probably do set vocabulary learning in spite of the objections to it from the currently popular audio-visual courses. A pilot experiment showed that the keyword technique, adapted for a group setting, improved both the immediate and delayed recall of 11-year-olds, while textual analysis suggested that the method might be applied to a useful portion of the words which children actually have to learn. A second experiment suggested that poor and successful learners approach vocabulary learning rather differently when left to their own devices, and further experimental groups confirmed the value of the technique for children of various levels of ability.

INTRODUCTION

RECENT American research has suggested a highly effective mnemonic technique which can be applied to the learning of foreign language vocabulary with dramatic results. The first published study in the area appears to be that of Ott *et al.* (1973), but most of the work since then has been carried out at Stanford University by Atkinson and his associates (e.g., Atkinson, 1975). The technique, developed from previous research on imagery and mnemonics, relates each foreign word to its English equivalent in two stages, the first being the formation of an 'acoustic link' between the foreign word and an English one with at least a partly similar sound. Thus the French word 'lapin' might suggest the English word 'lap', and this word becomes what Atkinson calls the 'keyword'. The next step is to form a vivid image relating the keyword to the actual meaning of the word in English—one might for example imagine a large white rabbit sitting on one's lap, so that a link has now been formed between the French word and its English equivalent.

Ridiculous though the technique may seem, it produces highly significant results. When subjects given keyword instructions are compared with control group subjects left to their own devices, keyword subjects frequently score 50 per cent better, and in one study Raugh and Atkinson (1975) report an experimental group score of 88 per cent in contrast to a control group score of 28 per cent. The Stanford research has looked more closely at certain aspects of the technique like experimenter-provided *versus* subject-generated acoustic and image links, and has shown that it can be applied successfully to both immediate and delayed recall of vocabulary in various languages. However, all these experiments involved student subjects, and only quite recently has the method been tried out with children. Pressley (1977) showed clearly that children, helped by interactive pictures, were able to use keyword method to improve their learning of Spanish vocabulary, and Pressley and Levin (1978) found age differences in children's ability to benefit from the technique. Several questions do remain unanswered, however, and, more important, the technique needs to be related much more closely to the actual demands made on children in the classroom. This study attempts to combine 'pure' experimental work based on the American research with a small-scale 'applied' study, to look at four areas:

- (1) The various audio-visual courses in use in many schools today tend to minimise the use of native language and the learning of individual words out of context, so do teachers actually want their pupils to learn lists of vocabulary items anyway? If not, the keyword technique may be psychologically

interesting, but of little direct educational value. Similarly, if teachers really do expect children to perform this task, what strategies are already being taught? In order to look at the status of vocabulary learning and the actual classroom demands made on children, an exploratory survey of classroom practice was proposed.

- (2) The experimental conditions used so far could also be related more closely to the classroom. The published studies on children have always used individual presentation, and Pressley and Levin (1978) consider that more evidence is needed to evaluate the technique in group settings. Equally important, they considered only immediate recall, and there is no published evidence that the technique facilitates children's long-term learning. Finally, recall of the foreign language word when given its English equivalent needs to be tested. All three points could be important if the technique is to be of practical use in schools, and a pilot study was proposed to examine these three questions.
- (3) The keyword method is clearly not suitable for all new words, and all the previous experiments have used carefully chosen items, selected for their probable appropriateness. (The studies on children, for example, have used only one or two syllable concrete nouns.) It is again fair to ask how widely the technique could be used in practice with words that children are actually expected to learn, and, to look at this point, an analysis of text of a widely used course was proposed. If the keyword technique proved suitable for only a tiny fraction of the words to be learned, its educational importance could again be questioned.
- (4) Finally, an experiment was designed to look in more detail at the value of the method for children of differing ability. For example, it is possible that almost any form of instruction could improve some children's vocabulary learning, and little is known about the strategies children spontaneously use on such a task when left to their own devices. The growth of comprehensive education has meant that almost all 11-year-olds are now expected to learn a foreign language, so that the question of the value of the technique for children of different levels of ability could be a vital one.

METHODS: PART I

Exploratory survey

Given that the survey was to be of an exploratory nature only, it was felt that a short questionnaire given to a number of children from different schools would be adequate. Lessons could have been observed, or teachers questioned, but it is possible that vocabulary learning is not the sort of activity teachers are proud of, and it was felt more important to assess the children's perceptions of their lessons rather than to look at what teachers say when questioned or do when observed. For example, teachers who set vocabulary learning may well also give detailed instructions on how it should be learned, but the question is whether or not children perceive, understand and follow such instructions. It was therefore also considered vital not to administer the questionnaire in school, where the classroom setting and presence of the teacher might well affect the answers, and both the pilot and main questionnaires were administered to children meeting in small groups (youth clubs, Scouts and Guides) outside the classroom and not in school time. While such children may well not be typical in some ways, it should be emphasised that the questions avoided matters of opinion and attitude and that the questionnaire was intended only as an exploratory survey of an area which previous studies of the keyword technique have tended to ignore. The topic would clearly be suitable for study in its own right, and the results presented here should be seen only as a first step in linking the experimental research with classroom practice.

A pilot version was first designed to look at four areas:

- (1) Roughly what proportion of children were expected to learn lists of vocabulary regularly, or occasionally?
- (2) How was such learning carried out—how many words were given, were the children tested, and how much direct help was given in methods of learning the items?
- (3) Since audio-lingual and audio-visual courses do not encourage such tasks, roughly what proportion of the children were taught at least partly by these methods, and were they still expected to learn vocabulary?
- (4) The original intention had been to ask in detail about any learning strategies proposed by teachers or used by pupils, but the pilot version suggested that this would be better covered by the use of actual items in the experimental study to be described later.

After some discussion, a pilot version of 11 questions was given to 27 girls and 13 boys, drawn from several schools, and the final version was modified in the light of their answers, with nine questions being given to a small number of children from 39 different schools. In all, 155 children, almost all in the 11-13 age range, answered the questionnaire. Each question and all the possible answers were read aloud to the children, who were assured of confidentiality and asked not to write down their names or those of their teachers. It was also emphasised that there were no right or wrong answers, and the children were told to cover up each question as soon as they had responded, to prevent copying, and to ask about anything they did not understand. Since it was clearly the most common foreign language the children learned, questions were restricted to French only. All the questionnaires were administered in April or May, so that the children had at least six months' experience on which they could draw.

RESULTS: PART I

The first question established that 87 per cent of the sample had actually done French during the previous months, and the rest of the analysis is obviously limited to their answers only. The proportion of children giving each possible answer is shown to the nearest 1 per cent.

In French lessons, did you ever have film strips or slides?

Often (say every few weeks or more often)	10%
Sometimes (say once or twice a term)	10%
Rarely (say once or twice ever)	18%
Never	63%

Although those with little experience may have found it hard to distinguish between some of the categories, the overall response is clear, and the vast majority of children reported that they rarely or never used the film strips or slides which are recommended as an important part of widely used audio-visual courses like the Longman's French course (Moore and Antrobus, 1973) or the Nuffield Introductory French course *En Avant* (1972).

Did you ever listen to tapes in French, or use a language laboratory?

(Answers defined as in the previous question)

Often	39%
Sometimes	21%
Rarely	23%
Never	17%

Use of tapes would therefore seem to be rather more common than the use of slides, for either practical or pedagogical reasons. When responses on these two questions were compared, a chi-squared test in fact suggested that teachers who did use slides were more likely to use tapes as well ($P < 0.05$).

Did you ever have to learn vocabulary? (This was defined for the children as 'lists of French words and what they mean in English', and response categories were as above.)

Often	67%
Sometimes	21%
Rarely	8%
Never	4%

This question was of particular interest, and the responses suggested very strongly that most children were quite often expected to learn lists of vocabulary. Chi-squared tests were again used to compare responses here with those to the previous two questions, but neither was significant at the 0.05 level. In other words, most children reported that their teachers set vocabulary learning at least sometimes, regardless of their use of slides or tapes.

If you did have to learn vocabulary, was it usually

Homework	28%
Classwork	27%
Both	45%

This question had been asked to see if vocabulary learning was perhaps regarded by teachers as the sort of task children could be left to do on their own, at home. In fact, the results are very evenly divided, suggesting that homework and classwork are equally common.

How many words, roughly, did you have to learn at a time? (Responses have been grouped for display purposes.)

Less than 10	10%
10-14	35%
15-19	18%
20-24	20%
Over 25	17%

The mean was 16.9 (sd 9.9), but the modal response was 10, with 26 per cent of the children giving that answer.

Were you usually given a test on the words?

Yes	80%
No	20%

Again, the traditional vocabulary test still seems common, though it has no place at all in the audio-visual courses, where vocabulary is supposed to be built up and tested indirectly, in context.

Did the teacher ever give you advice about how to learn vocabulary—what to do to help you remember the words? (Some examples were suggested here, like "Repeat the words over and over again" or "see if the French word reminds you of an English word".)

Often	8%
Sometimes	27%
Rarely	34%
Never	30%

It would thus appear that, although vocabulary learning was commonly set, the children felt that teachers did not often give detailed guidance as to how it should be learned.

Reliability. It might well be argued that children's perceptions of their lessons and their responses to questionnaires are not very reliable, and an attempt was made to examine this point by asking them at the end of the session to indicate if another person in the group was in the same French class at school. Of the 135 who did French, 41 were in the same class as at least one other person, and their papers were marked accordingly before being taken in. Ideally, their answers should have been identical on all seven questions shown above, but in practice there were 38 occasions when answers did not agree. (When three or more people were in the same class, a 'majority decision' was taken. For example, if two children put that they 'often' had slides and a third put that they 'sometimes' did, this was counted as one difference.) Further examination showed that disagreements on questions offering four choices tended to involve differences of only one category (only eight times did a child put an answer differing by two or three categories), and it also seemed that disagreements were evenly spread over all the questions, with no one clearly worse than the others. Agreement was thus reached 249 times out of 287, or 87 per cent of the time, suggesting that the answers were at least reasonably reliable.

DISCUSSION: PART I

In spite of its exploratory nature only, the questionnaire therefore does strongly suggest that many teachers still expect their pupils to learn and be tested on lists of 10–20 vocabulary items, often with little guidance. It might still be argued that such tasks are undesirable, however, and many language specialists have written opposing the use of native language and the learning of lists of individual words. The parallel with native language acquisition is often drawn (e.g., Mueller, 1974), and the pupil is seen as actively formulating hypotheses and learning new items in meaningful context with the aid of actions, objects and pictures, so that the use of English can be entirely avoided (e.g., Lambert, 1963). On the other hand, although such an approach has encouraged interesting materials and lively teaching methods, it tends to emphasise content and presentation while almost entirely ignoring the strategies actually used by learners—an area of growing interest for many psychologists (Ott *et al.*, 1976). The parallel between an infant acquiring native language and an 11-year-old learning French at school is highly misleading, and the social and cognitive differences between the two are immense. It is clear that language and concepts develop together, whatever view one takes of the exact relationship between them; the 11-year-old beginning a second language already has a large store of concepts and strategies inextricably tied in with his native language, and it is pointless to expect this language not to intrude, especially if teachers are dealing, as they are advised to, with ideas and situations that are already familiar (e.g., Preibusch and Zander, 1971). The use of pictures is a common method of trying to avoid native language, even though learners may well covertly label them anyway, perhaps even inappropriately—Hammerly (1974), for example, found that when subjects were presented only with a new Russian or German word and an appropriate picture, about 70 per cent of them immediately thought of a native language label which varied from about 75 per cent correct in the case of 'bird' to only 40 per cent in the case of 'deer'.

Learning lists of vocabulary should obviously not entirely replace the stimulating use of actual objects or pictures and the presentation of new words in context, but it would therefore appear that a case can be made for the overt use of English, at least in the early stages of learning a foreign language, and whatever theoretical arguments may be proposed, it also appears that in practice teachers still do frequently set vocabulary learning. Moreover, it is possible that at least some of the opposition to

the task was a result of growing dissatisfaction with dreary and ineffective rote learning (e.g., Rivers, 1968), so that if psychological research can suggest an effective alternative, then experimental studies like those involving the keyword method can be seen not only as psychologically interesting, but also as educationally relevant.

METHODS: PART II

The first experiment was a pilot study to see if the keyword technique could be adapted in a group setting to improve children's delayed as well as immediate recall and the learning of French words when given their English meaning as well as vice versa. It involved the comparison of two similar groups of children on two tests, given immediately after learning and again one week later. One group of children was taught by the keyword method and the other group learned the words in any way they wished.

Sample and materials

The subjects were 42 11-year-olds, all of whom had been learning French for about six months. They were all taught French by the same teacher, and were randomly assigned to the two groups. The experimenter and two colleagues selected a list of twelve French words with suitable keywords and English equivalents, and a check was made with the teacher to ensure that the children did not already know any of the items. Following the advice of Raugh and Atkinson (1975), the keywords were all monosyllables, and were similar to the first syllable of the French word wherever possible.

TABLE 1
VOCABULARY ITEMS USED IN PILOT EXPERIMENT

French	Keyword	Translation
lunettes	(net)	spectacles
tapisserie	(tap)	tapestry
imprévu	(pray)	unexpected
matelot	(mat)	sailor
loupe	(loop)	magnifying-glass
canapé	(can)	settee
lapin	(lap)	rabbit
coteau	(cot)	hillside
baguette	(bag)	rod
batelier	(bat)	ferryman
pamplemousse	(moose)	grapefruit
pataud	(pat)	clumsy person

Procedure

One group, chosen at random, served as the control group, and the other as the experimental or keyword group. Both were told that they would be tested on all 12 words, and both were given the same three practice items first. For the control group, each French word and its English equivalent appeared together on a large card. The experimenter first exposed the French word, pronouncing it twice, then the English word, which was also pronounced twice. Subjects were then instructed to learn the pair in any way they wished. After 10 seconds, the experimenter pronounced the pair again before going on to the next item. The experimental group had the same treatment except that after exposing and pronouncing each pair twice, the experimenter exposed and pronounced the keyword and used the ten seconds to describe the image link. For example:

Lunettes . . . spectacles . . . net

Imagine a football net with a pair of spectacles caught up in it, left there by the goalkeeper.

It should be noted that this variation of the technique therefore suggested an interacting image for the subjects without the need to provide an actual picture. All subjects were then given two tests:

Test 1—French to English

The experimenter exposed the 12 French words, previously written in random order on the blackboard, and pronounced each one in turn. Subjects simply wrote down what they thought the word meant in English, having been reassured that spelling would not count. There was no time limit.

Test 2—English to French

The published research has not really examined this aspect of vocabulary learning, but it was felt that a test requiring children to write down French words after a single exposure would be rather difficult, and so a recognition test was used, bearing in mind that results on the two tests could therefore not be compared directly. The 12 English words were read out in a different random order, and subjects were instructed to copy from the board the French word which they thought meant the same.

All available subjects were unexpectedly given the same two tests exactly one week later. Finally, control group subjects were asked informally about what strategies they had used, for possible use in the main experiment.

RESULTS: PART II

Table 2 shows the mean score out of 12 for each group on each test.

TABLE 2
MEAN SCORES BY CONDITION, PILOT EXPERIMENT

	Test 1	Test 2	Test 3	Test 4
Control	7.1	7.0	6.0	6.3
Keyword	8.7	9.3	9.1	8.3

There appeared to be a strong ceiling effect even on the delayed tests, particularly in the keyword group, where eight out of 20 subjects obtained perfect scores on the two immediate tests, and five obtained perfect scores on all four tests, and so non-parametric tests of significance were used. Mann Whitney U tests showed that subjects in the keyword group had recalled significantly more than control group subjects on all four tests ($P < 0.05$ in each case). A chi-squared test also indicated that there were significantly more maximum scores overall in the keyword group than in the control group ($\chi^2 = 7.0$, $P < 0.01$).

DISCUSSION: PART II

The pilot study thus supports Pressley's (1977) findings and suggests that the keyword method can be successfully used in a group setting, even without provided pictures, by 11-year-olds and not just college students, if both the keyword and a description of the image link are provided. The superiority of the experimental group is not so great as that frequently reported in the Stanford experiments, but this could be at least partly due to the ceiling effect. The pilot study also suggests that the technique might be particularly useful in children's delayed recall—obviously vital

in classroom learning—since overall the experimental group showed a superiority of 28 per cent over the control group on the immediate tests, and 42 per cent on the delayed tests. Finally, the results suggest that the keyword method also improves recall of foreign language items, given English equivalents, though the two tests cannot be directly compared and Test 1 may itself have had a rehearsal effect facilitating scores on Test 2, a point emphasised by Yuille (1973).

The study was, however, obviously limited, and raised several issues to be examined further, in the main experiment.

METHODS: PART III

Sample and design

The subjects in the second experiment were 194 11- and 12-year-olds, the whole first year intake of a local comprehensive school. There were 100 boys and 94 girls, and scores on the Schonell B reading test suggested a wide range of ability. They had been learning French for about eight weeks when the study was carried out.

The children were tested in randomly assigned groups, each group randomly assigned to one of the five conditions described below. Each child was given four tests, two immediately after learning and the same two tests exactly two weeks later.

Materials

All the children had been using the Longman's French course (Moore and Antrobus, 1973) and all the words were selected from the actual course materials, using a section which none of the children had yet reached. Analysis of text isolated the next 50 new words that the children would have to learn and, in view of the ceiling effect in the pilot study using 12 items, 18 of these words were selected as possibly suitable for the keyword method. Discussion with the teachers led to the rejection of one item and several modifications to the keywords chosen, reflecting Atkinson's advice that the choice of items is best left to a group of people rather than an individual. It should also be noted that some of the 50 French words (e.g., 'football', 'intelligent') were so similar to their English equivalents as to need no mediating link, so that, on the basis of this admittedly very small sample, it would seem that the keyword method might be suitable for a reasonable portion—at least a third—of the words which children are actually expected to learn.

TABLE 3
VOCABULARY ITEMS USED IN MAIN EXPERIMENT

French	Keyword	Translation
tard	(tar)	late
attraper	(a trap)	to catch
fatigué	(fat)	tired
lancer	(lance)	to throw
sage	(sergeant)	good
manger	(manager)	to eat
beaucoup	(bow)	many
porter	(porter)	to carry
tasse	(ass)	cup
épicerie	(a piece)	grocer's
tomber	(Tom)	to fall
travailler	(travel)	to work
trouver	(trove)	to find
sonnerie	(sun)	bell
ami	(me)	friend
matin	(mat)	morning
vilain	(villain)	naughty

Procedure

The purpose of the experiment—‘to look at different ways of learning new words in French’—was briefly described to all the children, and they were warned that they would be tested immediately after learning the 17 new words. No warning of the delayed tests was given. The five conditions were as follows:

‘*Test only*’. It was possible that subjects might already know, or be able to guess at the meanings of some of the words, particularly in the recognition test, so a small group of 19 children was given the two tests without any teaching at all, to establish a base-line against which other results could be judged.

Control Group ($N = 51$). This condition was very similar to that used in the pilot study, except that an overhead projector was used to keep presentation the same for all groups. The words were pronounced twice as the experimenter pointed to them on the screen, and the children were to learn each pair in any way they wished, followed by a final pronunciation. Thirty seconds were allowed for each item, and there were two practice items, both of which the children already knew (e.g., ‘*armoire*—cupboard’).

Image Group ($N = 51$). After a short discussion of forming ‘pictures in your mind’, the two practice items were used to instruct subjects. After pronouncing and pointing to the two words twice, as in the control condition, the experimenter then pointed to the keyword printed below, and described an interactive image to link the keyword to the English word. For example, the keyword for ‘*armoire*—cupboard’ was ‘arm’, and subjects were instructed to “imagine a cupboard on the wall. Suddenly the door bursts open and a big hairy arm comes out.” A final repetition of each pair filled up the remaining time of the 30 seconds allowed for each item.

Picture Group ($N = 49$). This treatment was identical to that of the Image group except that the description of the image was accompanied by the exposure of a simple line drawing.

Rote Group. Informal reports of children in the pilot study had suggested that rote repetition was by far the most common method spontaneously used by subjects in the control group, and so a small group of 24 subjects was actually instructed to use rote repetition in the main experiment, since it could be argued that any form of instruction whatsoever might increase scores. It is also possible that, in the keyword conditions, merely hearing the words repeated during the description of the image could have had a facilitative effect, so care was taken to repeat each pair four times before getting the children to repeat it quietly to themselves. As with all other groups, a final repetition of each pair ended the 30 seconds allowed for each item.

Testing procedure was exactly the same as in the pilot study, with a recall test of French into English (Test 1) being followed by a recognition test of English into French (Test 2). In addition, all subjects were asked to rate the method they had used on a simple five-point scale for enjoyment, ranging from ‘Liked it a lot’ to ‘Hated it’. Exactly two weeks later, the same two tests were given unexpectedly to all subjects except those in the ‘Test only’ condition, as Test 3 (French into English) and Test 4 (English into French).

An attempt was also made to discover what strategies had been spontaneously used by control group subjects. Informal discussion after the pilot study, and expectations based on the research literature, suggested five categories. Using the two practice items as examples, the experimenter slowly went through each pair with the control group, asking them to write down the label of method or methods they had used, along with any comments.

English. “The French word reminded you of an English word, and you used

this to help you remember the French word.” (This strategy is clearly related to the ‘acoustic link’ of the keyword method.)

Image. “You made up a picture in your mind to help you to remember the word.” (This strategy could be partly related to the ‘image link’ of the keyword method. It was called ‘picture’ for the benefit of the children, but will be referred to as ‘image’ to avoid confusion with the use of external pictures.)

Repeat. “You repeated the two words several times to yourself.” (This strategy had been very popular in the pilot study and is obviously similar to the Rote condition.)

Other. “You used some other way of learning the words.” (Subjects were encouraged to give details, including the possibility that they already knew the word.)

Don't know. It was emphasised that this category did not imply low ability, and that subjects simply might not know or remember how they had learned the words.

It was possible that some children might have found this rating task difficult, and that those who remembered most words would also be more likely to remember what strategies they had used. In fact, neither of these seems to have happened, as discussion of the results will show.

RESULTS: PART III

Taking each condition separately, highly significant correlations were obtained between scores on all four tests ($P < 0.003$ in all cases), suggesting that scores could be combined if necessary on the various tests. Analysis of variance showed that there were no significant differences in the reading ages of the five groups, suggesting that any differences between their test scores could not be due to this factor. ($F = 0.36$, $df = 4$, $P < 0.05$.)

Table 4 shows the mean score of each group on each test, out of a maximum possible of 17.

TABLE 4
MEAN SCORES AND ENJOYMENT RATINGS BY CONDITION, MAIN EXPERIMENT

	Test 1	Test 2	Test 3	Test 4	Enjoyment Rating
Control	7.1	9.1	5.9	7.6	3.2
Rote	9.5	10.3	7.0	9.2	2.7
Image	12.6	13.4	11.2	12.8	3.7
Picture	13.7	14.4	12.4	14.0	4.2
Test only	0.5	2.4	—	—	—

Taking each test separately, analysis of variance showed that there were highly significant differences between the various groups' scores, at well beyond the 0.01 level in each case. Separate *t*-tests showed that not all the differences were significant, however. The difference between Rote and Control groups was barely significant on immediate recall, $t(76) = 1.96$, $P < 0.05$, one-tail test only, and was not significant on the delayed test, $t(72) = 1.03$, $P > 0.05$. Similarly, the Picture group's score was not significantly higher than the Image group's on immediate recall, $t(90) = 1.2$, $P > 0.05$. The ‘test only’ group did very badly, suggesting that subjects were rarely able to guess at the answers or use previous knowledge, and the results of the keyword groups were clearly much better than those of the control and rote repetition groups.

Analysis of variance suggested that enjoyment ratings had very little direct effect, explaining less than 0.2 per cent of the total variance above and beyond that caused

by the other factors ($F = 0.1$, $P > 0.05$). Sex and reading age both had significant effects on the total variance—sex about 1 per cent and reading age about 9 per cent, while the various conditions had a much larger effect, explaining almost 26 per cent of the total. The interaction between condition and sex was not significant, but the interaction between condition and reading age was, so the relationship was examined in more detail by splitting the reading age scores into three groups of roughly equal size, as in Table 5, which shows the mean total out of 68 for each group.

TABLE 5
MEAN TOTAL SCORES BY CONDITION AND READING AGE, MAIN EXPERIMENT

	RA Under 10	RA 10.0-11.0	RA Over 11
Control	15.3	30.1	39.4
Rote	32.5	34.8	48.8
Image	39.8	47.6	57.8
Picture	46.5	48.5	64.4

Amongst other things, these results suggest that even rote instructions brought dramatic improvements to the scores of children in the low RA group, while the provision of a picture along with keyword instructions enabled them to treble the scores they would have obtained when left to their own devices, and to reach a level higher than that of control group subjects in the high RA group.

Finally, an attempt was made to analyse the reported strategies of the control group to see if they were related to success in the tests. Taking the group as a whole, it is clear from Table 6 that some spontaneous strategies were more common than others, and that they differed in the success they brought on the recall tests. Although the children had been encouraged to put down more than one strategy if necessary,

TABLE 6
SPONTANEOUS STRATEGIES REPORTED BY CONTROL GROUP SUBJECTS, MAIN EXPERIMENT

	English	Image	Repeat	Other	Don't Know
% Frequency reported, all group	28	4	36	5	22
% Success on recall tests, all group	56	59	30	72	25
% Frequency reported, high scorers	49	5	22	11	18
% Frequency reported, low scorers	14	2	62	5	21

such reports were in fact very rare, and occurred only about 3 per cent of the time. (Interestingly, the combination of 'Image' and 'English', which contains both elements of the keyword method, gave a success rate of 88 per cent, though it was reported only 1 per cent of the time.) Rote repetition, while the most common strategy chosen, was much less successful than the 'English' strategy. Imagery also gave good results, but was reported very rarely, while the high scores in the unclassified ('other') category were explained by analysis of the children's responses, which showed that in most cases they had already known the word.

Taking total scores on all four tests for control group subjects, a high-scoring group and a low-scoring group, the top and bottom 25 per cent respectively, were

isolated. *t* tests showed large differences between their scores (52.8 and 11.4) and reading ages (12.6 and 9.8), both significant at well beyond the 0.01 level, and there were also some striking differences between the reported strategies of the two sub-groups, as shown in Table 6. A chi-squared test showed first that the two sub-groups did not differ significantly on their reports of 'don't know' what strategy they used ($P > 0.05$), suggesting that both groups were able to report equally successfully on how they had learned the words. Neither group had made much use of imagery, and the difference here was again not significant at the 0.05 level, but further chi-squared tests showed that they did differ significantly ($P < 0.01$ in each case) on their reported use of rote repetition and the 'English' strategy. The low-scoring subjects spontaneously used rote repetition much more often than the high-scoring subjects, who were frequently able to see similarities between the French word and English words for themselves, and these results coupled with those of the experimental groups suggest that the use of different strategies is a very important factor in explaining differences between high and low scores.

DISCUSSION

Taken together, the questionnaire, pilot study and main experiment strongly support the value of the keyword method in children's classroom learning of French vocabulary. Although there have been frequent objections to the task as part of a foreign language curriculum, the questionnaire suggested that most teachers still use vocabulary learning and testing regardless of what other techniques they also use, though the children did feel that they were often given very little help as to how the learning should be done. These findings offer a practical justification for studies of vocabulary learning, but the task can also be defended on theoretical grounds because many learners still appear to use covert native language equivalents even if the teacher tries to avoid them at all costs.

Rote repetition is not the only way of learning such material, and the pilot study suggested that 11-year-olds are quite capable of using the keyword method, modified for group presentation, to learn the English meanings of French words and vice versa, for both immediate tests and tests given unexpectedly a week later.

In the main experiment, preliminary analysis of text suggested that the method could probably be applied to a useful proportion of the words which children are actually expected to learn, and the keyword technique gave results much better than those obtained by rote learning instructions or by leaving children to their own devices, as they probably would have been in the classroom. Pressley and Levin (1978) found that pictures were not essential for 11-year-olds to benefit from the technique, although the present study suggests that they may be useful for less able children. However, there was still a ceiling effect even with 17 items, and over a quarter of the subjects in the 'Picture' group obtained perfect scores on all four tests, a feat equalled only by two subjects in the Image group and one in the Control group. The questionnaire had suggested that 17 words were the mean number set in practice, so that the extra work involved in producing pictures might not bring much better results, but might enable teachers to increase the number of words set, this being a possible area for further research.

One problem with redintegrative processes like mnemonics is the danger of interference, but examination of the errors made by keyword subjects showed that intrusions were rare—only occasionally did a child confuse the keyword and the translation, with one exception—the 'false friend' 'travailler', which was translated by a number of subjects as 'travel'. However, examination of the other groups showed that this particular mistake was equally common there, suggesting that it was not necessarily interference from the keyword which had caused the error. Another

form of interference was the occasional semantic intrusion from the image link rather than the keyword. For example, given 'tomber, Tom Thumb, to fall' and an image of a tiny Tom Thumb walking across the desk and tripping over your thumb, a few children put 'little' or 'small' as the meaning of 'tomber'. However, both kinds of interference were rare, and the research suggests that it becomes less likely as time goes on and a direct link is formed between the foreign word and its equivalent (Atkinson, 1975).

A more fundamental objection to the keyword method might be the arbitrary and artificial nature of the mediating links it forms. Admittedly, it would be better to point to existing semantic links where they exist, but such links may not always be apparent to the children who are expected to learn the items. 'Fatigué', for example, might seem obviously related to the English word, but when the 19 children in the 'Test only' group were presented with the list of 17 French words and asked to guess which one meant 'tired', only two guessed correctly, and one of these admitted that she already knew it.

The keyword method is not being advocated as suitable for all new material, or even as equally helpful for all children. It is not intended to replace the presentation of new items in a variety of contexts or the use of other methods, but can be seen as another technique, and apparently an enjoyable and highly effective one, to be added to the teacher's existing repertoire. If much of what we expect children to learn in school in fact appears highly arbitrary (Bower, 1973), then mnemonics, which are particularly useful in the learning of arbitrary material, might have an even wider role to play, and there is perhaps room for more research which tries to relate experimental findings to actual classroom materials and demands, combining the approaches of pure and applied research. Certainly, the present study strongly supports previous research on the keyword mnemonic and suggests that it can be effectively applied to children's classroom learning by giving artificial 'meaning' and by extending the principles of meaningful learning to material which could otherwise be learned only by rote.

NOTE.—The data collection for the pilot study was carried out in collaboration with J. French.

REFERENCES

- ATKINSON, R. C. (1975). Mnemotechnics in second language learning. *Am. Psychol.*, 30, 821-828.
- BOWER, G. H. (1973). Educational applications of mnemonic devices. In DOYLE, K. O. (Ed.), *Interactions: Readings in Human Psychology*. Lexington, Mass.: D. C. Heath.
- HAMMERLY, H. (1974). Primary and secondary associations with visual aids as semantic conveyers. *I.R.A.L.*, 12, 118-125.
- LAMBERT, W. E. (1963). Psychological approaches to the study of language. Part 2: on second language learning and bi-lingualism. *Modern Language J.*, 47, 114-121.
- MOORE, S., and ANTROBUS, A. L. (1973). *Longman Audio-Visual French*, Book 1. London: Longman.
- MUELLER, T. (1974). Another look at how to teach listening and reading comprehension. *Modern Language J.*, 58, 19-23.
- NUFFIELD FOREIGN LANGUAGES PROJECT (1972). *En Avant*. Leeds: E. J. Arnold.
- OTT, C. E., BLAKE, R. S., and BUTLER, D. C. (1976). Implications of mental elaboration for the acquisition of foreign language vocabulary. *I.R.A.L.*, 14, 37-48.
- OTT, C. E., BUTLER, D. C., BLAKE, R. S., and BALL, J. P. (1973). The effect of interactive-image elaboration on the acquisition of foreign language vocabulary. *Language Learning*, 23, 197-206.
- PREIBUSCH, W., and ZANDER, H. (1971). Vocabulary teaching: a search for an analytical model. *I.R.A.L.*, 9, 131-145.
- PRESSLEY, M. (1977). Children's use of the keyword method to learn simple Spanish vocabulary words. *J. educ. Psychol.*, 69, 465-672.
- PRESSLEY, M., and LEVIN, J. R. (1978). Developmental constraints associated with children's use of the keyword method of foreign language vocabulary learning. *J. exper. Child Psychol.*, 26, 359-372.

CAUSAL PERCEPTIONS OF SUCCESS AND FAILURE BY ADVANTAGED, INTEGRATED AND DISADVANTAGED PUPILS

BY A. RAVIV, D. BAR-TAL,* ALONA RAVIV AND Y. BAR-TAL
(*Tel-Aviv University, Israel*)

SUMMARY. One hundred and thirty-four sixth-grade pupils belonging to different social groups, upper middle class of European-American and Israeli origin (advantaged pupils), lower middle class of European-American, Asian-African and Israeli origin (integrated pupils), and lower class of Asian-African origin only (disadvantaged pupils) were asked to attribute causality for their success or failure, following a real test in mathematics. The results showed that in the case of success all three groups of pupils tended to attribute the outcome more to internal than to external causes and more to stable than to unstable causes. In the case of failure, advantaged pupils tended to attribute the outcome more to internal than to external causes, integrated and disadvantaged pupils tended to attribute their failure more to stable rather than unstable causes. The implications of these results are discussed.

INTRODUCTION

WEINER and his associates (Weiner *et al.*, 1971; Weiner, 1974) have suggested that causal perceptions of success and failure may be of major importance in understanding achievement-related behaviour. In order to explain achievement-related behaviour, they proposed an attributional model based on the assumption that causal perceptions of success and failure mediate between the antecedents of causal perception and achievement performance.

Specifically, they suggested that beliefs regarding the causes of success and failure can be classified on two dimensions: locus of control and stability. Taking as examples four different causes (ability, effort expended, task difficulty, and luck), the classification is performed in the following manner: (a) while ability and effort are considered internal because they originate within the person, task difficulty and luck are external because they rest outside the person; (b) while ability and task difficulty are considered stable owing to the fact that they do not vary in the event of the same task being reattempted, effort and luck are considered unstable because they may vary over time.

These dimensions of locus of control and stability are important, respectively, in the understanding of affective reactions of pride and shame to success or failure and the changes in perceived probability of success for future outcomes (see Weiner, 1974). Thus, for example, individuals feel maximum pride in a success situation, and maximum shame in a failure situation when the attributions are internal, since personal responsibility is assumed. Also, attributions of failure to unstable causes (especially internal, such as lack of effort), result in higher expectancies for future success than attributions pertaining to stable causes—the supposition being that it would be possible to improve the performance if more effort were to be invested. Failure attributed to stable causes, such as lack of ability, results in low expectancy for future success, since it is believed that stable causes do not change over time.

Furthermore, the types of causes individuals utilise to explain their successes or failures are important determinants of their achievement-related behaviour. For example, individuals who tend to attribute their failure to stable causes, such as lack of ability, tend not to persevere in the face of failure, since they believe that ability

*Requests for reprints should be sent to: Daniel Bar-Tal, School of Education, Tel-Aviv University, Tel-Aviv, Israel.

is a stable element which does not allow for the possibility of future change in the outcome. On the other hand, those pupils who tend to attribute their failure to unstable-internal causes, such as lack of effort, tend to persist in the event of failure, since they believe effort to be a factor which is changeable and allows for the possibility of future modification of the outcome. (See Weiner, 1974, 1977, and Bar-Tal, 1978, for a review of empirical evidence.)

The effect of causal perception on achievement-related behaviour has important implications, in view of evidence that the pattern of causal perceptions might differ in various social groups. In this vein, Coleman *et al.* (1966) and Katz (1967) suggested that the cognitive processes pertinent to achievement motivation may be learned differentially by various ethnic and socio-economic groupings. For example, Katz suggested that certain ethnic groups may not develop cognitive structures supporting the efficacy of effort, i.e., they do not perceive the covariation between effort and outcome.

Attributional research has relatively neglected the investigation of causal perception of success and failure of different socio-economic groups. A few experiments comparing attributions of socio-economic class have been performed in a laboratory setting, but the results of these studies were not conclusive. For example, Friend and Neale (1972) studied attributional patterns of 120 fifth-grade children who were equally divided into four groups on the basis of race and social class. The children were instructed to read a short paragraph and subsequently their memory was tested. The outcomes were manipulated and, following the experimenter's feedback, the children were asked to attribute their success or failure to four causes of ability, effort, task difficulty, and luck. The results of this study indicated that regardless of socio-economic level, white children judged ability and effort (internal causes) to be more important than task difficulty and luck (external causes), especially following failure feedback. Black children, on the contrary, judged task difficulty and luck to be more important causes than ability and effort. In another study, Falbo (1975) instructed 48 five-year-old children to listen to a taped story about a child assembling a puzzle, and asked them to ascribe causality either to the successful or unsuccessful completion of the task. The results of this study showed that middle-class children tended to choose effort as the explanation of an outcome more often than lower-class children.

The attributional patterns of various social groups can be inferred indirectly from research dealing with a comparison of attributions between individuals with high and low need for achievement. A number of studies have shown that individuals with high need for achievement tend to attribute success to ability and effort (internal causes) and failure especially to lack of effort (internal-unstable cause). In addition, they perceive themselves as having high ability. Individuals with low need for achievement tend to attribute success to various causes, without any clear preference, and attribute failure especially to lack of ability which is considered an internal-stable cause (see Weiner, 1974, for review). McClelland (1961) suggested that children of lower classes (disadvantaged children) are characterised by low achievement motivation. Thus, it is possible to suppose that the attributional pattern of disadvantaged children is similar to that of individuals with a low need for achievement. The socialisation patterns of middle and upper middle classes encourage the development of high achievement motivation (McClelland, 1961), and, therefore, it is possible to suppose that the attributions of advantaged children are similar to those of individuals with a high need for achievement.

The present study compared attributions of pupils of three different social groups, regarding academic performance in actual classroom settings. Possible attributions were not limited to four causes, but were extended to nine empirically-derived ones (see Bar-Tal and Darom, 1979). The first group of subjects consisted of advantaged pupils from the upper middle class, mostly of European-American

and Israeli origin, the second group consisted of integrated pupils from the lower middle class of European-American, Asian-North African and Israeli origin and the third group consisted of disadvantaged pupils from the lower class of Asian-North African origin only. (The term 'integrated' is used to express the fact that the population of the investigated township consists of people of Asian-African and European-American origin who send their children to the same schools).

The hypotheses for the present study were as follows:

- (a) Advantaged pupils would tend to attribute their success more to internal than to external causes, and more to stable than to unstable causes. They would tend to attribute failure more to internal than to external causes—favouring internal-unstable causes in particular.
- (b) Disadvantaged pupils would not display clear attributional preferences for success, but failure, on the other hand, would be attributed more to internal causes rather than external ones—especially to internal-stable causes.
- (c) Finally, we expected that the attributional pattern of the integrated pupils would share features of both the advantaged and disadvantaged pupils and would therefore not be clearly definable.

METHOD

The sample

The sample comprised 134 sixth-grade pupils (66 boys, 68 girls) drawn from two classes in each of three different schools—each school being located in a distinct socio-economic area. The first group of pupils, whom we termed 'advantaged', (24 boys, 27 girls) were from a school located in a predominantly upper middle socio-economic level neighbourhood in northern Tel-Aviv. The residents of this area were mainly European-American or Israeli in origin, as opposed to Asian-African.* The second group (29 boys and 27 girls) was termed the 'integrated' group. They came from a lower middle class township east of Tel-Aviv where the population showed a balance in origins between European-American and Asian-African.† The third group (13 boys and 14 girls) came from a school designated as 'disadvantaged'††, located in a township east of Tel-Aviv with a mainly lower class, Asian-African population. The characteristics of the three groups of pupils are presented in Table 1. We note that, moving from 'advantaged' through 'integrated' to 'disadvantaged' pupils, there is an *increase* in the average number of household members, and a *decrease* in the average number of rooms per household, mean level of father's education, and mean IQ level.

Procedure

Following a final, trimester examination in mathematics, the experimenter (the last author) was contacted by the teacher prior to the return of the graded tests. After the graded tests had been returned, the experimenter came to the classroom and distributed a questionnaire. The instructions were typed on the first page of the questionnaire and were read aloud by the experimenter. They read as follows:

"The purpose of the questionnaire you have just received is to examine pupils' thoughts regarding examinations in school. In the following pages, you'll be asked questions regarding your feelings and thoughts about the examination and grade which you have just received. In this questionnaire, there are neither right nor wrong answers. Please be honest when answering the questions and write

* This ethnic classification is commonly used with reference to the Jewish population by the Israeli Census Bureau Abstracts and in research studies.

† The comparisons of integrated pupils according to their origin reveal that they do not differ with regard to the variables appearing in Table 1

†† The Israeli Ministry of Education classifies schools as disadvantaged on the basis of established formal criteria (see Algarebali, 1975).

TABLE 1
CHARACTERISTICS OF THE THREE GROUPS

Variables (N)	Advantaged School (51)	Integrated School (56)	Disadvantaged School (27)
Origin*			
Asia-Africa	7.8%	42.9%	100%
Europe-America	42.2%	44.6%	—
Israel	50.0%	12.5%	—
Average number of people in a household	4.5	5.0	7.2
Average number of rooms in a household	3.9	3.3	3.1
Mean father's education†	5.1	3.0	1.9
Mean IQ	116.6	106.0	93.5

* The origin of the pupil was determined on the basis of the father's place of birth.

† The father's education was determined on a 6-point scale: (1) didn't complete elementary school; (2) completed elementary school; (3) completed junior high school; (4) completed high school; (5) post-high school education; (6) university degree.

what you really feel or think. Don't skip any questions. The answers to these questions will not be shown to your teachers" (English translation). Following the reading of the instructions, the pupils were asked to answer the questionnaire.

Questionnaire

The questionnaire consisted of two parts. First, the subjects were asked to evaluate the received grade (the 'outcome') as a success or failure. Then the subjects were asked to indicate the relative influence of each of nine listed causes on the level of the grade received. The following list of causes, derived from a study by Bar-Tal and Darom (1979), was used: ability in mathematics, interest in mathematics, difficulty of the material, effort exerted during the test, difficulty of the test, preparation for the test at home, teacher's explanation of the material, learning conditions at home and luck. The answers were given on a 7-point scale from (1) indicating 'no influence' to (7) 'very great influence'.

RESULTS

Several analyses were carried out to compare the responses of the subjects. Since no sex differences were detected on any of the variables, the data of both sexes were combined.

Causal perception

Comparisons within the groups. The mean attributions for each cause according to group of pupils and outcome are presented in Table 2. Figure 1 describes graphically the success and failure attributional patterns of each group of subjects. In Figure 1, we see that all three groups of pupils tended to attribute success mainly to the teacher's explanation, personal interest in the subject matter, personal ability. Differences among the three groups are more evident in the case of failure. The advantaged pupils tended to attribute failure mainly to deficiencies in effort exerted during the test, teacher's explanation, and preparation for the test at home. The integrated pupils tended to attribute their failures fairly equally to a number of causes: low levels of teacher's explanation, low effort invested, and low level of

TABLE 2
MEAN ATTRIBUTIONS

Group of Pupils	Outcome	Causes								
		Ability	Interest	Material difficulty	Effort	Test difficulty	Preparation	Teacher	Home Conditions	Luck
Advantaged Pupils	Success (N=37)	4.59	4.95	3.08	4.21	2.97	3.24	5.74	3.18	2.31
	Failure (N=14)	2.29	2.36	1.79	3.00	2.07	2.71	2.79	1.29	2.36
Integrated Pupils	Success (N=41)	5.17	5.12	3.05	4.56	2.68	3.49	5.22	4.24	2.61
	Failure (N=15)	3.72	3.61	3.82	3.89	3.89	3.50	3.94	2.94	1.89
Disadvantaged Pupils	Success (N=17)	5.50	4.94	3.78	4.11	4.44	3.61	4.71	4.17	3.22
	Failure (N=10)	4.75	5.00	4.92	4.17	4.55	4.64	5.17	4.00	3.50

personal ability and interest, and high levels of test and material difficulty. Disadvantaged pupils tended to attribute their failures mainly to teacher's poor explanation, low personal interest and ability, and the difficulty of the material.

In order to perform comparisons between the groupings of causes within the three groups, two independent judges classified the nine causes according to the two dimensions of locus of control and stability, using the Elig and Frieze (1975) coding scheme. With reference to locus of control, ability, interest, effort, and preparation were classified as internal causes and difficulty of the material, test difficulty, teacher's explanation, home conditions, and luck were classified as external causes. With regard to stability, effort, preparation, and luck were categorised as unstable causes, and the remaining six were categorised as stable. Table 3 shows the mean attribution levels for the types of causes and the results of statistical comparisons. The following findings were obtained: (a) Advantaged pupils tended to make higher attributions to internal rather than to external causes and to stable rather than to unstable causes in the case of success, and in the case of failure, they tended to make higher attributions to internal than to external causes. (b) Integrated and disadvantaged pupils tended to make higher attributions to internal rather than to external causes in the case of success, and, for both success and failure, higher attributions were made to stable causes as opposed to unstable ones.

Comparisons among the groups. In order to compare the attributional patterns of the three groups of pupils, a Profile Analysis (see Morrison, 1967) was performed for the attributions of the three groups of pupils in situations of success and of failure (i.e., six groups in total). The comparison was carried out with multivariate analysis of variance on the eight differences between the degree of causal ascription of each two consequent causes. (Nine causes were presented in the same order for all the groups.) The analysis yielded a significant largest root criterion of 0.211, $P < 0.05$; the parameters are $s = 5$, $m = 1$, $n = 59.5$ —indicating that the attributional patterns are not similar in all the groups.

Causal Perception of Pupils

FIGURE 1
MEAN ATTRIBUTIONS OF THE NINE CAUSES IN SITUATIONS OF SUCCESS AND FAILURE BY THE THREE GROUPS OF PUPILS

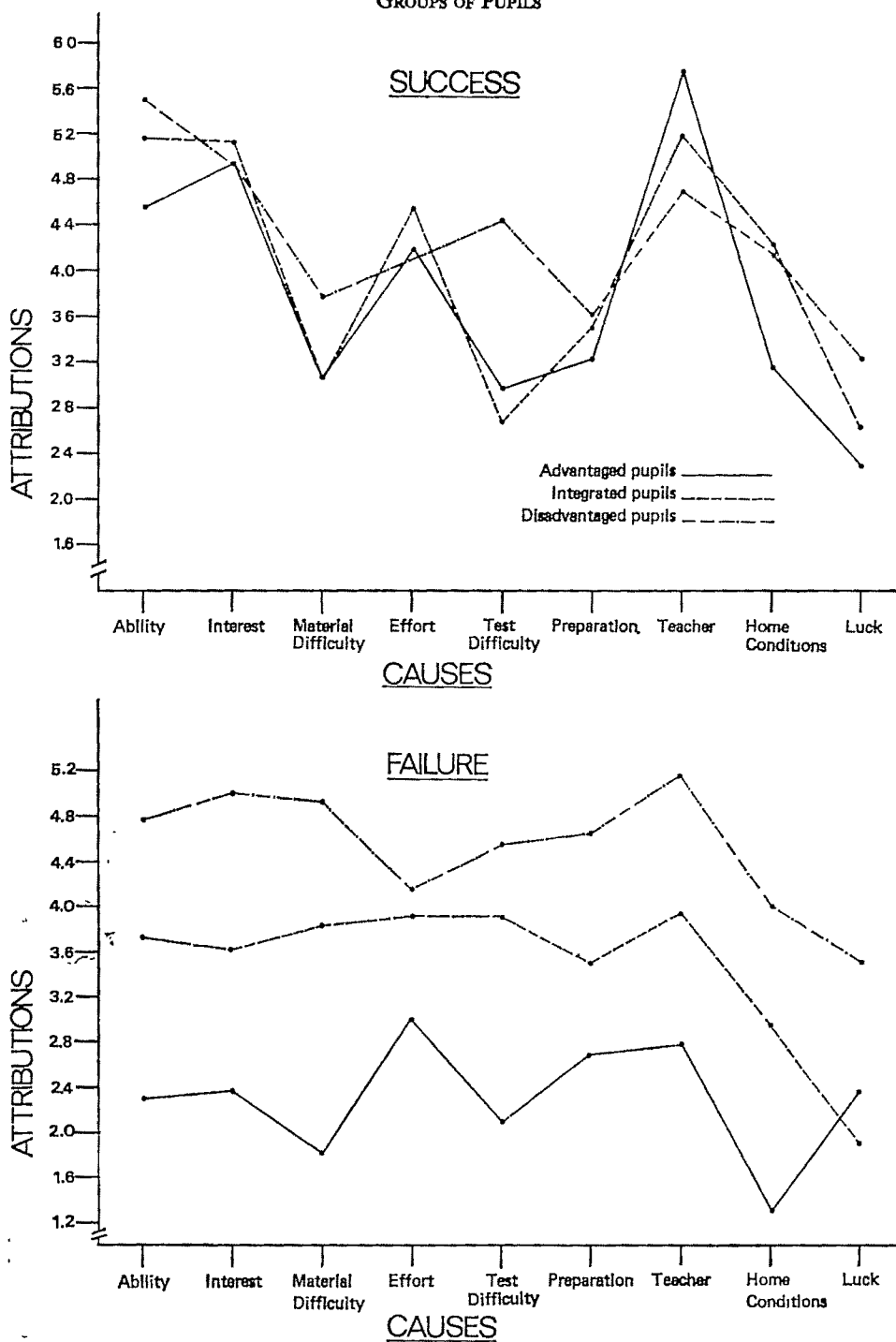


TABLE 3

MEANS AND T-TESTS BETWEEN TYPES OF CAUSES ACCORDING TO GROUPS OF PUPILS AND OUTCOMES

Group of Pupils	Outcome (N)	Type of Cause	Mean	Type of Cause	Mean	df	t
Advantaged	Success (N=37)	Internal	4.26	External	3.43	36	3.48**
		Stable	4.10	Unstable	3.22	36	3.41**
	Failure (N=14)	Internal	2.59	External	2.06	13	1.95*
		Stable	2.33	Unstable	2.70	13	1.02
		Internal-stable	2.32	Internal-unstable	2.85	13	1.12
Integrated	Success (N=41)	Internal	4.58	External	3.56	40	4.92**
		Stable	4.12	Unstable	3.55	40	3.84**
	Failure (N=15)	Internal	3.60	External	3.25	14	1.06
		Stable	3.73	Unstable	2.91	14	2.42*
		Internal-stable	3.67	Internal-unstable	3.53	14	.27
Disadvantaged	Success (N=17)	Internal	4.51	External	4.03	16	1.98*
		Stable	4.47	Unstable	3.70	16	3.63**
	Failure (N=10)	Internal	4.52	External	4.36	9	.41
		Stable	4.77	Unstable	3.90	9	2.26*
		Internal-stable	4.80	Internal-unstable	4.25	9	.96

* $P < 0.05$, one-tailed test** $P < 0.01$, one-tailed test

In addition, univariate one-way analyses of variance were performed for the six groups on the difference between the degrees of causal ascriptions on each of two consequent causes. The results, presented in Table 4, which indicate that there are differences in attributional patterns among the six groups, made it necessary to analyse the attribution for each cause separately. The nine causes were thus treated as dependent variables in nine successive 3×2 analyses of variance (Group of pupils $\times$ Outcome). The results of these analyses (summarised in Table 5) yielded main effects for Outcome on ability, interest, effort, teacher's explanation, and home conditions. That is, pupils who succeeded believed that these above-mentioned causes were more influential in the achievement of the outcome than did pupils who failed. Main effects for the Group of pupils on ability, difficulty of material, test difficulty, and home conditions indicate that while disadvantaged pupils considered these causes as the most influential in the outcome, advantaged children considered them to be the least influential. Finally, four two-way interactions were obtained: (a) while the advan-

TABLE 4

F VALUES OF THE UNIVARIATE ANALYSES OF VARIANCE ON THE DIFFERENCES BETWEEN THE CAUSAL ASCRIPTION OF TWO CONSEQUENT CAUSES

Differences between the Two Causes	df	F
Ability—Interest	5,128	0.78
Interest—Material difficulty	5,128	2.39*
Material difficulty—Effort	5,128	2.31*
Effort—Test difficulty	5,128	3.95**
Test difficulty—Preparation	5,128	1.21
Preparation—Teacher	5,128	2.70*
Teacher—Home conditions	5,128	2.47*
Home conditions—Luck	5,128	2.10

* $P < 0.05$ ** $P < 0.01$

TABLE 5
ANALYSES OF VARIANCE SUMMARY FOR THE NINE CAUSES

Source	df	Ability	Interest	Material diffi- culty	Effort	Test diffi- culty	Prepara- tion	Teacher	Home Condi- tions	Luck
Group of Pupils (A)	2,130	5.43**	2.06	4.69*	.63	7.37**	1.86	.07	6.79**	2.53
Outcome (B)	1,130	23.46**	18.62**	.09	4.17*	.41	.00	15.00**	9.91**	.45
A × B	2,130	1.79	4.41*	3.89*	.65	3.78*	.79	6.13**	1.47	.36

* $P < 0.05$ ** $P < 0.01$

taged and integrated groups of pupils considered interest to be a more influential cause in the case of success than in the case of failure, the disadvantaged group of pupils considered interest to be equally influential in both situations; (b) while the advantaged group of pupils considered material difficulty to be a more influential cause in the case of success than in the case of failure, the integrated and disadvantaged groups of pupils thought the opposite; (c) while the advantaged group of pupils considered test difficulty to be a more influential cause in the case of success than in the case of failure, the integrated group of pupils displayed the opposite tendency, and the disadvantaged group of pupils evaluated test difficulty similarly in both cases of success or failure; and (d) while the advantaged and integrated groups of pupils considered the teacher's explanation to be a more influential cause in the case of success than in the case of failure, the disadvantaged group of pupils displayed an opposite tendency.

DISCUSSION

The results of this study were found only partially to support our original hypotheses that (i) the clearest attributional preferences would be that advantaged students would attribute success to internal-stable causes and failure to internal-unstable causes, (ii) the clearest preference of disadvantaged students would be to attribute failure to internal-stable causes, and (iii) the middle integrated group would show no clear preferences, sharing tendencies of the other groups. The results did not show the clear distinction we would have expected with regard to success, but there were clear distinctions in the patterns for attribution of failure. We feel, however, that the findings indicate that attributional patterns do characterise social groups.

Specifically, the findings with regard to the comparisons within each group indicated that in a situation of success, all three groups of pupils (not just the advantaged group) tended to make higher attributions to internal rather than external causes and to stable rather than unstable causes. In a situation of failure, advantaged pupils did tend to make higher attributions to internal rather than external causes, while both disadvantaged and integrated pupils tended to make higher attributions to stable rather than unstable causes. These results resemble the findings regarding the attributional differences between individuals with high- and low-need for achievement which indicated that the main difference between the two groups exists in the case of failure, while in the case of success, the difference was not consistent (see Weiner, 1974). Furthermore, although the comparison within groups did not show a difference between the attribution of integrated and disadvantaged pupils, further comparative analyses among the three groups (discussed below) revealed that the attributional patterns of integrated children does fall somewhat in between the patterns of advantaged and disadvantaged children. However, it does resemble more the pattern of disadvantaged than advantaged pupils. The direct comparisons among the attributions of the three groups for cases of success and of failure (see Table 5) revealed

some interesting differences regarding the attribution of *stable* causes. First, disadvantaged pupils perceived ability and home conditions to be, in general, more influential causes than did either advantaged or integrated pupils perceive. Secondly, while advantaged pupils perceived interest, difficulty of material, test difficulty, and teacher's explanation (all stable causes) to be more influential on the outcome in the case of success than in the case of failure, disadvantaged pupils perceived difficulty of material and teacher's explanation to be more influential on the outcome in the case of failure than in the case of success, and interest and test difficulty to be equally influential—both in cases of success and failure. These results indicate that advantaged pupils, more so than disadvantaged ones, exhibit a tendency to attribute success to stable causes, while disadvantaged pupils tend to attribute failing more to stable causes than do advantaged pupils.

If we apply the analysis of the affective consequences of attributions (Weiner, 1974) to the attributional patterns for success shown by three groups, the inference is that all the pupils experienced pride as a result of their success, since they assumed personal responsibility for success. Moreover, the successful pupils could expect repetition of their success, since their tendency was to attribute it to stable causes (see Weiner *et al.*, 1976). The latter tendency was found, according to the direct comparisons among the three groups, to be especially profound among the advantaged pupils. In the case of failure, the advantaged pupils experienced shame because of their tendency to make internal attributions. However, they attributed failure similarly to both stable and unstable causes, with a minor tendency towards unstable ones (see Figure 1). Thus, although they may experience shame, they still can hope for a change in a future outcome. Such a hope may greatly influence their achievement behaviour in the direction of persisting in the face of failure. In contrast, integrated and disadvantaged pupils would seem not to persist in the future, but rather to display a behaviour of helplessness. Their tendency to attribute failure to stable causes means that no change in outcome can be expected in the future, and that, therefore, failure may be repeated (see Dweck, 1975).

It could be thought that the resemblance of the attributions for failure on the part of integrated and disadvantaged children is a result of the fact that the failing pupils in the integrated classes were of Asian-African origin. However, examination of the data indicates that, out of the 15 failing integrated pupils, six were of European-American origin and their attributions did not differ from the other nine failing pupils of Asian-African origin. The resemblance is probably due to the attachment to socio-economic class rather than to different ethnic origins. It seems that the attributional patterns of pupils from the lower middle socio-economic class is more similar to pupils from the lower class than to pupils from the upper middle class.

In conclusion, we would like to note the usefulness of the study of children's causal perception in explaining differences in achievement-related behaviour. One specific point should be stressed. The detrimental pattern of causal-attribution for failure, which disregards the importance of unstable causes, need not be a hopeless situation: recent evidence suggests that it is possible with the aid of training programmes to modify maladaptive attributions which may, in turn, subsequently affect achievement behaviour and improve academic achievement. This task rests, to a large extent, on teachers who can often substantially influence the attributional patterns of their pupils (see Bar-Tal, 1979).

ACKNOWLEDGMENTS.—This research was supported by a Grant-in-Aid for Research provided by the School of Education, Tel-Aviv University. The authors are grateful to the pupils and teachers who agreed to participate in the study. The authors also thank Ilana Gelernter for her help in analysing the data and Alex Ingersoll for his comments on the earlier draft of this paper.

REFERENCES

- ALGAREBALL, M. (1975). Indices for the characterization of the social composition of schools and a system for the allocation of budgets to schools with disadvantaged pupils. *Megamot*, 21, 217-227 (Hebrew).
- BAR-TAL, D. (1978). Attributional analysis of achievement-related behavior. *Rev. Educ. Res.* 48, 259-271.
- BAR-TAL, D. (1979). Interactions of teachers and pupils. In FRIEZE, I. H., BAR-TAL, D., and CARROLL, J. S. (Eds.). *New Approaches to Social Problems: Applications of Attribution Theory*. San Francisco: Jossey-Bass.
- BAR-TAL, D., and DAROM, E. (1979). Pupils' attributions of success and failure. *Child Dev.*, 50, 264-267.
- COLEMAN, J. S., CAMPBELL, E. Q., HOBSON, C. J., MCPARTLAND, J., MOOD, A. M., WEINFELD, F. D., and YORK, R. L. (1966). *Equality of Educational Opportunity*. Washington, D.C.: United States Department of Health, Education, and Welfare.
- DWECK, C. S. (1975). The role of expectations and attributions in the alleviation of learned helplessness. *J. Pers. Soc. Psychol.*, 31, 674-685.
- ELIG, T., and FRIEZE, I. (1975). A multi-dimensional scheme for coding and interpreting perceived causality for success and failure events: The CSFS. *JSAS Catalog of Selected Documents in Psychology*, (Ms. No. 1069), 5, 313.
- FALBO, T. (1975). The achievement attributions of kindergarteners. *Dev. Psychol.*, 11, 529-530.
- FRIEND, R. M., and NEALE, J. M. (1972). Children's perceptions of success and failure: an attributional analysis of the effects of race and social class. *Dev. Psychol.*, 7, 124-128.
- KATZ, I. (1967). The socialization of academic achievement in minority group children. In LEVINE, D. (Ed.), *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.
- MCCLELLAND, D. C. (1961). *The Achieving Society*. New York: Free Press.
- MORRISON, D. F. (1967). *Multivariate Statistical Methods*. New York: McGraw-Hill.
- WEINER, B. (1974). *Achievement Motivation and Attribution Theory*. Morristown, N.J.: General Learning Press.
- WEINER, B. (1977). An attributional approach for educational psychology. In SHULMAN, L. (Ed.), *Review of Research in Education*. Itasca, Ill.: F. E. Peacock.
- WEINER, B., FRIEZE, I., KUKLA, A., REED, L., REST, S., and ROSENBAUM, R. M. (1971). *Perceiving the Causes of Success and Failure*. New York: General Learning Press.
- WEINER, B., NIERENBERG, R., and GOLDSTEIN, M. (1976). Social learning (locus of control) versus attributional (causal stability) interpretations of expectancy of success. *J. Personality*, 44, 52-68.

(Manuscript received 14th May, 1979)

THE DEVELOPMENT OF A CO-OPERATIVE, COMPETITIVE, AND INDIVIDUALISED LEARNING PREFERENCE SCALE FOR STUDENTS

By L. OWENS

(University of Sydney)

AND R. G. STRATON

(Murdoch University)

SUMMARY. An important variable in the effectiveness of learning is the preference of the student for a mode of learning, i.e., co-operative, competitive, or individualised. A Learning Preference Scale is therefore an asset, both to teachers as they try to suit the mode of classroom activity to student preferences, and to researchers as they investigate the dynamics of learning, especially aptitude-treatment interaction. The *Learning Preference Scale—Students (LPSS)* was developed with a sample of 1,643 Sydney school-children. Internal consistency, test-retest stability, sub-scale inter-correlations, and factor analysis show that the LPSS has considerable promise as a valid and reliable instrument. Results show clear and significant differences between the preferences of boys and girls over the range of grades from Year 4 to Year 11.

INTRODUCTION

IN recent years, much attention has been focused upon the classroom use of co-operative, competitive, and individualised learning modes or 'goal structures' (Johnson and Johnson, 1974). Traditionally, classroom practice has emphasised the competitive learning mode with students encouraged to work alone and to compete with each other on tests and examinations for the rewards of the education system. However, this competitive mode has increasingly attracted strong criticism from advocates of co-operative and individualised learning.

The co-operative learning mode, exemplified historically in Dewey's Project Method and employed in Bruner's social science curriculum, "Man: A Course of Study", is much in evidence in the current practice of innovative and open plan schools. The individualised learning mode, on the other hand, was given considerable impetus two decades ago with the development of programmed instruction. This has been followed by several elaborate, integrated schemes such as Individually Prescribed Instruction, Individually Guided Education, and the Personalised System of Instruction. Although some research effort has been devoted to techniques of organising for co-operation and individualisation in the classroom, only very recently has interest been shown in student attitudes towards the various learning modes, and the attitudinal and achievement correlates of these preferences.

Definitions

As Johnson and Johnson (1975) have defined it, a learning mode or goal structure specifies:

"the type of interdependence existing among students and the ways in which students will relate to each other and to the teacher in working toward the accomplishment of instructional goals" (p. 7).

Relying on the seminal research of Deutsch (1949a, 1949b), they identify three different states of interdependence:

- (a) A co-operative goal structure exists when students can achieve their own individual goals only by working conjointly with others as they achieve their goals (e.g., a large task accomplished by division of labour);

- (b) A competitive goal structure exists when students can achieve their own individual goals only when others fail to achieve their goals (e.g., coming first in a test or a race);
- (c) An individualistic goal structure exists when students can achieve their own individual goals no matter what others have chosen to do (e.g., completing an individually-assigned library research assignment).

In this paper, the term 'learning mode' is used to refer to the state of interdependence that exists among students and teachers and to their use of materials and space in learning situations.

Review of research

Initially, reviews of research linking co-operative, competitive, and individualised goal structures with learning outcomes grouped the studies broadly as being concerned with 'process variables', 'cognitive outcomes', and 'affective outcomes' (Johnson and Johnson, 1974, 1975). With many more studies available, recent reviews (Johnson and Johnson, 1978, 1979) make a much more differentiated analysis of learning outcomes and are accompanied by increasingly confident claims about the relative merits of the three learning modes. Student attitudes toward, or approval for, each of these learning modes figure prominently in the conclusions reached about the effectiveness and satisfaction that students find in working together and alone. Thus, the time is apposite to assess the methods used to determine these attitudes.

In a recent general statement of their position, Johnson and Johnson (1978) have identified ten categories of 'instructional outcomes' that can be linked with the three learning modes. When these learning modes have actually been employed, they have been shown to affect four types of instructional outcome: learning complex material, ability to take the perspective of others, handling controversy and disagreement, and liking for types of teaching method. In addition, student attitudes to these learning modes have been shown to be related to six types of affective or attitudinal variables: motivation to learn, involvement in learning activities, attitudes towards peers, attitudes towards school staff, self-esteem and psychological health. Clearly, in these latter cases the method of obtaining student attitudes to the alternative learning modes is of critical importance.

There are four studies which are cited repeatedly as evidence for relationships in the six affective or attitudinal outcome areas, and they contribute significantly to the conclusions reached in each case (Johnson and Ahlgren, 1976; Johnson and Norem-Hebeisen, 1977; Norem-Hebeisen and Johnson, 1977; Johnson *et al.*, 1978). With regard to the area of psychological health, however, only Johnson and Norem-Hebeisen (1977) is cited. In each of the four studies, attitudes to co-operation, competition, and individualisation are assessed either from sub-scale scores on the Minnesota School Affect Assessment (Johnson, 1974) or from scores computed from the Social Interdependence Scales (Johnson and Norem-Hebeisen, 1979), which is a self-report inventory adapted from the Minnesota School Affect Assessment. Thus, the adequacy of each of these instruments needs to be examined.

The Minnesota School Affect Assessment (MSAA) is composed of 157 items, many of which comprise multiple rating scales, yielding a total of 240 separate responses. There are 35 clusters of items, ranging from 2 to 20 items per cluster, from which sub-scale scores are computed. Each of the 'Co-operation', 'Competition', and 'Independence' clusters has three items, and it seems reasonable to question the adequacy of three responses as the basis of conclusions about a student's attitudes. In addition, there appear to be inconsistencies in the way the items are phrased. For instance, although all three co-operation items refer to 'other students',

two of the three competition items refer to 'my friends', a much smaller frame of reference. Furthermore, the only negatively phrased item among the nine appears in the independence cluster. An additional peculiarity of this item, 'I don't like to work in groups', is that it is more an anti-group statement than it is pro-independence. In sum, these considerations fail to create confidence in conclusions based only on MSAA data (Johnson and Ahlgren, 1976; Johnson *et al.*, 1978).

The adaptation and expansion of these MSAA sub-scales into the Social Interdependence Scales has overcome some of the faults but has created others. In their final form, after a series of trials (Johnson and Norem-Hebeisen, 1979), the Co-operative and Individualistic Interdependence Scales each have seven items, and the Competitive Interdependence Scale has eight, a total of 22 items. The increased number of items at least improves the reliability of the scales when compared with the MSAA, and possibly their face validity as well. Anomalies in the placement of negatively phrased and reverse-scored items persist in the new scales, however. There are no such items in the Co-operative Scale, one in the Competitive Scale, and four in the Individualistic Scale, making the latter a scale with a different impact on the reader than the other two. More serious than this, however, is the repetitiveness of the items in two of the scales. In the Co-operative Scale the items are usefully differentiated across a wide range of types. In the Competitive Scale seven of the eight items are based only on statements of global preference, as are five of the seven items in the Individualistic Scale. If the aim is to obtain an expression of attitudes that is not influenced by the questionnaire itself, one doubts that this set of scales is completely satisfactory.

In summary, Johnson and his associates have based their claims for certain attitudinal outcomes largely upon correlations between students' self-reported attitudes towards the three learning modes and other attitudinal variables. To the extent that such claims involve attitudinal data gathered by means of the MSAA or the Social Interdependence Scales one is entitled to some scepticism. This is not to deny that there is a considerable weight of evidence to justify these claims, and that some of this evidence is derived from robust experimental studies. It is to say, however, that the portion of the evidence which is based upon relationships derived from the use of either of the two named instruments merits further examination. It is apparent, then, that in this area of considerable importance in classroom process research, an area with major theoretical and practical significance, there is a need for a valid and usable scale to measure student attitudes and preferences for the three learning modes. Such an instrument could also be used to test the sufficiency of the earlier results of Johnson and his associates. The bulk of this paper is an account of the development and testing of a scale which attempts to fill this gap.

METHOD

The Learning Preference Scale for Students: pilot testing

An initial version of the Learning Preference Scale: Students (Form A) was developed and pilot tested in Years 4, 6, and 9 in one primary and one secondary school. The wording of the instructions and the timing of the administration were tried out with whole classes. In addition, students from Years 5, 8, and 11 at the same schools were taken aside in groups of three and 'talked through' the items in the scale. These trials and subsequent data analyses led to the development of a revised version of the Learning Preference Scale: Students, Form B, which was used subsequently in the main trial (Barnes *et al.*, 1977).

Description of the instrument

Form B of the Learning Preference Scale: Students (LPSS) contains 42 items. Each item is a brief statement about a feature of learning by co-operating with others,

by competing with others, or by working alone. Jackson has suggested that 'substantial construct validity' is directly related to the extent to which items "are derived from an explicitly formulated, theoretically based definition of a trait" (1971, p. 232). Accordingly, a list was prepared of areas within which preferences for a learning mode might be expressed. This was further refined into 14 categories and an item related to each learning mode was developed within each category. The 14 items concerned with a particular learning mode became a sub-scale of the LPSS. Thirteen of these categories were content matched while the last category contained three significant but unmatched items. A description of the content categories is given below.

- (1) Global Preference: an overall preference for working together with a group, working on one's own, or striving to be best.
- (2) Global Dislike: an overall dislike for working in a particular mode, as given above, i.e., these items are similar to those in the Global Preference category but are phrased negatively.
- (3) Global Projection: a supposition about what 'other students' prefer, based on the psychoanalytic concept of projection.
- (4) Positive Work Outcomes: indicating that good quality outcomes will follow from working in a particular mode.
- (5) Negative Work Outcomes: indicating that poor quality outcomes will follow from working in a particular mode, i.e., similar to category 4 but phrased negatively.
- (6) Source of Ideas: indicating that the best ideas come from individuals working on their own or in combination with others.
- (7) Tension: indicating that personal stress is caused by working in a particular mode. These items are phrased negatively.
- (8) Altruism: indicating a belief that working in a particular mode is of benefit to others.
- (9) Reward: indicating that recognition of effort is clear and satisfying when working in a particular mode.
- (10) Rate of Progress: indicating that the work gets done quickly in a particular mode.
- (11) Decision Making: indicating that the quality of a decision is enhanced by working in a particular mode.
- (12) The Future: indicating that the experience of working in a specific mode is good preparation for post-school years.
- (13) Individual Differences: indicating that it is helpful to the teacher, in coming to know individual students, by arranging to work with them in a particular mode.
- (14) Unmatched: a set of items unrelated in content.

An example of the items comprising one of the categories may give a more complete impression of how the operational definition is carried into the actual phrasing of the LPSS. The three items in Positive Work Outcomes (with preference in brackets) are: 'the work of a group is usually better than if each member works alone' (co-operative), 'trying to be better than others makes me work well' (competitive), 'I do better work by myself' (individualised).

Students respond to each LPSS item by indicating how 'true' or 'false' the statement is for them. A four-point answer scale is used, and the response categories are described to students as 'completely true', 'a bit more true than false', 'a bit more false than true', and 'completely false'. An answer scale with an even number of choices was chosen to avoid the middle-of-the-range 'don't know' response.

Sample

One state high school and a neighbouring primary school were selected from each of four separate north-western or western suburbs of Sydney. Each high school was a full six-year co-educational institution, and ranged in size from 950 to 1,300 students. The areas were selected to provide a population with a broadly representative socio-economic background and with a relatively low number of non-English-speaking migrant residents. By the standard indicators of socio-economic status, the suburbs ranged from 'low' to 'upper middle' in social character (Davis and Spearritt, 1974; Connell *et al.*, 1975).

Testing was carried out on two occasions, the main testing ($N = 1643$) in August 1977, near the end of the second term of the school year, and the retesting ($N = 446$) in October in the middle of the third term (see Table 1). The main testing involved two intact classes per year in Year 7 to Year 11 in each high school and in Year 4 to Year 6 in each primary school. In the secondary schools, classes generally were chosen from the upper two-thirds of the range of ability-streamed groups in school

TABLE 1

NUMBER OF BOYS AND GIRLS FROM EACH SCHOOL YEAR IN THE PRIMARY AND SECONDARY SCHOOLS IN THE SAMPLE

		Main testing School year									Retesting School year											
School	Sex	Primary			Secondary						Total	Primary			Secondary							Total
		4	5	6	7	8	9	10	11	4		5	6	7	8	9	10	11				
A	B	30	39	27	16	21	26	21	18	198	14	15			17		11		57			
	G	24	31	40	29	26	23	29	17	219	11	15			10		20		56			
B	B	24	20	30	23	28	26	20	8	179		8	17	13		12	12		62			
	G	39	26	43	25	22	20	31	32	238		12	19	11		8	17		67			
C	B	31	22	28	23	27	21	27	20	199			9		19	15		12	55			
	G	29	26	27	31	24	23	21	17	198			11		10	17		5	43			
D	B	26	25	24	29	36	25	21	15	201	10			14				21	45			
	G	34	43	9	30	28	15	26	26	211	20			18				23	61			
Boys		111	106	109	91	112	98	89	61	777	24	23	26	27	36	27	23	33	219			
Girls		126	126	119	115	100	81	107	92	866	31	27	30	29	20	25	37	28	227			
Total		237	232	228	206	212	179	196	153	1643	55	50	56	56	56	52	60	61	446			

subjects. Low ability remedial classes were specifically avoided. In the smaller primary schools there were only two classes per year. In the larger primary schools with ability-streamed classes the top two classes were tested. Where a school had parallel ability classes a random choice of classes was made. There were a few minor variations to this systematic sampling plan caused by the situation within specific schools.

The retesting was conducted in a reduced number of classes. From the four primary schools taken together two classes were required in each of the Years 4, 5, and 6 meaning that only a single class from either one or two of the three years was retested from each school. In the four secondary schools, a similar sampling decision meant that in some schools a single class from three years and in other schools from only two years was retested (see Table 1).

Administration

The LPSS was printed so as to allow for optical scanning of responses and each student was provided with a dark lead pencil for marking answers. The initial

instructions read to each class contained a brief general statement of who the testers were and what the test was about, how to mark responses in the spaces provided and two practice questions. There was no obvious confusion, not even with the younger primary children. The LPSS was part of a battery of three tests which were all being trialled during the testing programme. The other two tests were an Australian revision of the MSAA and a Classroom Learning Atmosphere Scale (for high schools only). At most, a pair of these three tests was given to any one class in the sample, and the order of presentation was counterbalanced among years and schools across the total sample. In each year overall, the same number of classes did the LPSS first in the sequence as did it later, and the same number of classes had LPSS paired with each of the other two tests.

Generally, one visit per school was sufficient, but in some cases a return visit was necessary in order to test students who were absent on the first testing day. In the primary schools, it was expected that some students in Year 4 and Year 5 would experience difficulty in reading the items. After consultation with class teachers, a small number of slow readers, and in some instances non-readers, were removed from some classes before testing commenced. In addition, a small number of uncompleted papers bearing obvious signs of non-comprehension were removed from the sample afterwards. It seems that in Year 4 the better readers understood all of the LPSS items while the remainder grasped the sense of most of them. In all years, students were encouraged to indicate problems of reading or understanding and these difficulties were recorded for later use in revising items.

RESULTS

The data gathered during the trial of the LPSS were analysed to yield information on the reliability and the validity of the scale and each of its three sub-scales. In addition, differences in sub-scale scores were examined for groups differing in sex and school year. The results of these analyses are presented below.

Reliability

Both internal consistency and stability coefficients were determined for each LPSS sub-scale and these are presented in Table 2. The Cronbach α coefficients calculated on both the main testing data and retest data indicate a substantial degree of homogeneity for each of the sub-scales. These coefficients vary from 0.64 to 0.81. The test-retest coefficients, on the other hand, indicate only a modest degree of stability over the two-month interval between testings. It appears from these results that between 21 and 45 per cent of the variation in scores on the retest can be accounted for by variation in the scores on the initial test. While it is possible that this apparent lack of congruence between the two sets of coefficients is due to sample differences, this is unlikely given the way the retest sample was chosen. What is more likely is that these preferences are somewhat unstable and may vary in relation to time of testing within a single school year, a possibility which is being investigated in more detail in current studies using the LPSS. Interestingly, there are no clear year level or sub-scale differences among the internal consistency or stability coefficients, except that preferences for the co-operative learning mode appear to become more stable with an increase in age.

Validity

Information concerning the validity of the LPSS was obtained through factor analyses as well as through intercorrelation of the *a priori* sub-scales within the LPSS and between the LPSS and equivalent MSAA (Australian Revision) sub-scales for the initial test data. The results of the sub-scale intercorrelations are presented in Table 2. These indicate that within the LPSS there is virtually no covariation

TABLE 2

CO-OPERATIVE, COMPETITIVE, AND INDIVIDUALISED PREFERENCE SUB-SCALE RELIABILITY AND STABILITY DATA, INTERCORRELATIONS FOR LPSS SUB-SCALES, AND CORRELATIONS WITH MSAA SUB-SCALES

		LPSS Sub-scales		
		Co-op.	Comp.	Ind.
		<i>Internal consistency (Cronbach α)</i>		
Main testing (N = 1643)	Overall	0.81	0.66	0.74
	Years 4-6 (N = 697)	0.67	0.64	0.68
	Years 7-9 (N = 597)	0.70	0.72	0.72
	Years 10-11 (N = 349)	0.71	0.71	0.64
	Overall	0.68	0.68	0.73
		<i>Stability (Pearson r)</i>		
Two-month interval (N = 421)	Overall	0.54	0.56	0.58
	Years 4-6 (N = 157)	0.46	0.57	0.57
	Years 7-9 (N = 158)	0.54	0.57	0.57
	Years 10-11 (N = 106)	0.67	0.54	0.62
		<i>Correlation coefficients</i>		
LPSS Sub-scales (N = 446)				
Co-operative	Overall		-0.07	-0.39
	Years 4-6		-0.09	-0.39
	Years 7-9		-0.09	-0.39
	Years 10-11		0.03	-0.38
	Overall			0.30
Competitive	Years 4-6			0.26
	Years 7-9			0.38
	Years 10-11			0.31
	Overall			0.31
MSAA Sub-scales (N = 733)				
Co-operative		0.53	-0.09	-0.26
	Competitive	-0.12	0.60	0.24
	Independence	-0.48	0.16	0.49

between the Co-operative and the Competitive sub-scales while the Individualised sub-scale shows a modest amount of covariation negatively with the Co-operative sub-scale and positively with the Competitive sub-scale. The results are very similar when computed separately for three sub-groups of the total sample, Years 4 to 6, Years 7 to 9, and Years 10 to 11. Correlations between the LPSS and the MSAA sub-scales show a similar pattern, with those between equivalent sub-scales being the most substantial, ranging from 0.49 for the Individualised sub-scales to 0.60 for the Competitive sub-scales.

Comparisons of the LPSS inter-sub-scale correlation coefficients computed separately for different year level groups show interesting differences between the present study and the findings of Johnson *et al.* (1978) using the MSAA sub-scales and of Johnson and Norem-Hebeisen (1979) using the Social Interdependence sub-scales (see Table 2). In the present study the Co-operative-Individualised correlations are virtually the same across different year levels, whereas in the previous studies

these correlations, while negative, increase in magnitude with year level. In addition, the Competitive-Individualised correlations in Johnson and Norem-Hebeisen's (1979) study are virtually zero for Year 12 and college students while being modest (0.22 and 0.32) for their two samples of Year 7 students. However, these coefficients are modest and comparable for all year levels in both the present study and that conducted by Johnson *et al.* (1978).

These findings lend support to Johnson and Johnson's (1975) notion, supported by their own research (Johnson and Ahlgren, 1976), that preferences for co-operative, competitive, and individualised learning modes comprise three independent preference dimensions. This is in contrast to earlier findings which showed preferences for co-operation and competition to be negatively related and opposed in a polar relationship. However, these earlier findings typically emerged from studies where a subject was forced to choose between co-operation and competition rather than being free to rate preferences for each learning mode independently, as is the case with the LPSS, and MSAA, and the Social Interdependence Scales.

It was noted above that the correlations between the LPSS and the MSAA sub-scales revealed a similar pattern to that found within the LPSS, with the highest relationship being shown by the negative correlations between the Individualised and the Co-operative sub-scales. This relationship within the LPSS seems largely to be due to two sets of items which require a preference to be expressed with respect to working in a group and to working alone. As might be expected, there was a tendency for subjects to express a high preference for one set of items and a low preference for the other set, even though they were not constrained to do so as was the case in some earlier research.

The factorial validity of the LPSS was determined by subjecting its 42 items to principal factor analysis using the SPSS programme FACTOR (Nie *et al.*, 1975). A nine-factor solution was obtained for the main testing data using the default parameter values specified in the programme. As this solution included four single item factors, a five-factor solution was obtained and subjected to both varimax and oblique rotations. Table 3 presents the results of the oblique rotation using a factor loading of 0.30 as the criterion for including an item in a factor while Table 4 compares the varimax and oblique rotations.

The results of these analyses are quite striking in that after oblique rotation each of the first three factors contains eight items which in each case come exclusively from a single *a priori* sub-scale, the Individualised, Competitive, and Co-operative

TABLE 3
PRINCIPAL FACTOR ANALYSIS OF LPSS ITEMS FOR MAIN TESTING DATA: FIVE SPECIFIED
FACTOR SOLUTION WITH OBLIQUE ROTATION (N = 1643)

Item no.	LPSS item	Sub-scale	Loading*
<i>Factor I</i>			
8	I do not like working by myself.†	I	0.62
18	I do better work by myself.	I	0.58
24	I work badly when I know I have to do it all by myself.†	I	0.57
16	If I work by myself most of the time, I become lonely and unhappy.†	I	0.53
4	I prefer to work by myself so I can go as fast as I like.	I	0.51
26	I like my work best if I do it myself without anyone's help.	I	0.43
23	If I work by myself now I will manage better later.	I	0.36
35	I like to work on my own without paying attention to other people.	I	0.32

TABLE 3 (*continued*)

Item no.	LPSS item	Sub-scale	Loading*		
<i>Factor II</i>					
31	Trying to be better than others makes me work well.	Cm	0.65		
42	I like to try to be better than other students.	Cm	0.61		
40	Trying to be better than others in schoolwork helps me be successful later.	Cm	0.56		
38	I do not like always trying to be better than someone else.†	Cm	0.49		
21	I learn faster if I'm trying to do better than others.	Cm	0.48		
22	I do not mind if I get the lowest marks.†	Cm	0.39		
36	My work is not so good when I'm thinking mostly about doing better than other people.†	Cm	0.38		
11	I would like to be the best at something.	Cm	0.38		
<i>Factor III</i>					
14	I like to work in a group at school.	Co	0.66		
30	Working in a group now helps me work with other people later.	Co	0.42		
7	Working in a group scares me.†	Co	0.41		
39	I do not like working with other people.†	Co	0.39		
19	I like to help other people do well in a group.	Co	0.36		
9	I like to be in a group which people know is good.	Co	0.35		
3	It is helpful that the teacher gets to know students by having them work together in groups.	Co	0.34		
17	We get the work done faster if we all work together.	Co	0.31		
<i>Factor IV</i>					
41	I make good decisions when I listen only to the advice I give myself.	I	0.53		
32	I like to keep my ideas to myself.	I	0.44		
35	I like to work on my own without paying attention to other people.	I	0.43		
12	Other people's ideas are not usually as good as mine.	Cm	0.38		
26	I like my work best if I do it myself without anyone's help.	I	0.31		
25	Other people get a lot out of trying to be better than me.	Cm	0.30		
<i>Factor V</i>					
15	I like to be able to use the ideas of other people as well as my own.	Co	0.39		
5	It is helpful to combine everyone's ideas when making a decision.	Co	0.35		
<i>Factor pattern correlations</i>					
	Factor I	Factor II	Factor III	Factor IV	Factor V
Factor I	1.00	0.19	-0.22	0.08	-0.00(4)
Factor II		1.00	-0.04	0.21	0.13
Factor III			1.00	-0.19	0.30
Factor IV				1.00	0.20
Factor V					1.00

* A criterion factor loading of 0.30 was used for including an item in a factor.

† This item has been reverse scored because it is phrased negatively.

sub-scales respectively. Factor IV contains four items from the Individualised sub-scale and two items from the Competitive sub-scale, while Factor V contains only two items, both from the Co-operative sub-scale. All factors are unipolar. The factor pattern correlation matrix, also presented in Table 3, indicates only a small degree of relationship between the factors after rotation to the oblique solution.

TABLE 4

RELATIONSHIP BETWEEN THE VARIMAX AND OBLIQUE ROTATIONS FOR THE FIVE FACTOR SOLUTION OF THE PRINCIPAL ANALYSIS OF THE LPSS ITEMS: MAIN TESTING DATA (N = 1643)

	Varimax rotation factors	Oblique rotation factors*				
		I	II	III	IV	V
I	+	4I, 8I, 16I 18I, 23I, 24I, 26I, 35I				
	-			14 Co, 17 Co		
II	+		11 Cm, 21 Cm, 22 Cm, 31 Cm, 36 Cm, 38 Cm, 40 Cm, 42 Cm			
	-					
III	+			3 Co, 7 Co, 9 Co, 14 Co, 17 Co, 19 Co, 30 Co, 39 Co		
	-					
IV	+				12 Cm, 25 Cm, 26I, 32I, 35I, 41I	
	-			39 Co		13 Co
V	+					5 Co, 15 Co
	-					

Other items: 1, 2, 6, 10, 20, 27, 28, 29, 33, 34, 37

* All of the factors from the oblique rotation were unipolar.

Three additional samples or sub-samples were also subjected to parallel sets of analyses each of which involved a principal factor analysis followed by both varimax and oblique rotations, five factors having been specified. The samples used were the 446 subjects who were involved in the retesting, and two sub-samples from the initial testing, i.e., the odd-numbered subjects (N = 832) and the even-numbered subjects (N = 811). Each of these additional analyses yielded solutions strikingly similar to that obtained using the complete main testing data for the first three factors, and also for the fourth factor in all but one instance. Taken together, the results of these analyses give strong empirical support to the *a priori* sub-scale structure of the LPSS.

Group differences

Mean scores were calculated for boys and girls separately at each year level using the main testing data (Table 5). These results have been standardised within each LPSS sub-scale and are presented graphically in Figure 1. The major feature of these graphs is the consistently different results for boys and girls despite the variations across the year levels. Girls expressed a greater preference for the co-operative learning mode than did boys, at each year level. On the other hand, boys expressed a greater preference for both competitive and individualised learning modes than did girls, again at each year level. Differences between the preferences of students at

TABLE 5
YEAR LEVEL $\times$ SEX MEAN SCORES AND STANDARD DEVIATIONS FOR THE
CO-OPERATIVE, COMPETITIVE, AND INDIVIDUALISED SUB-SCALES OF THE LPSS

Group	Co-operative		Sub-scale Competitive		Individualised	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Year 4	42.6	6.63	38.2	6.64	40.5	7.35
Boys	41.2	6.57	39.8	6.27	41.1	7.20
Girls	43.9	6.45	36.9	6.68	39.9	7.46
Year 5	43.7	5.80	38.3	6.54	40.0	6.59
Boys	43.0	5.21	39.5	6.06	40.8	6.09
Girls	44.2	6.22	37.3	6.77	39.4	6.94
Year 6	42.9	6.59	38.1	6.27	40.7	6.48
Boys	42.0	6.65	38.7	5.45	41.5	5.97
Girls	43.7	6.45	37.6	6.91	40.0	6.86
Year 7	44.0	5.95	40.7	6.72	39.8	6.98
Boys	43.0	5.86	42.5	5.80	40.8	6.93
Girls	44.7	5.95	39.3	7.10	39.0	6.97
Year 8	43.4	5.95	39.3	6.27	37.9	6.38
Boys	42.0	6.40	41.3	5.89	39.3	5.77
Girls	44.9	5.00	37.1	5.95	36.4	6.70
Year 9	41.9	6.26	40.5	6.61	39.6	6.51
Boys	40.9	5.83	42.5	5.95	40.6	6.65
Girls	43.2	6.56	38.0	6.61	38.6	6.21
Year 10	43.6	5.75	40.2	5.83	39.3	5.83
Boys	43.0	6.02	42.1	5.27	39.6	5.88
Girls	44.1	5.50	38.7	5.83	39.1	5.81
Year 11	42.6	5.71	40.0	6.31	39.3	5.08
Boys	41.7	5.59	41.4	6.42	40.3	4.83
Girls	43.1	5.74	39.1	6.10	38.7	5.16
Overall	43.1	6.13	39.3	6.48	39.7	6.53

different year levels were not so clear, however, although there appeared to be a tendency for competitive learning mode preferences to increase and for co-operative and individualised learning mode preferences to decrease with increases in year level. Two-way (Year Level $\times$ Sex) analyses of variance of these data, using $\alpha = 0.05$ for determining significance, revealed no significant interaction effects, but there were significant main effects for both Year Level and Sex for each sub-scale (Table 6). The rather erratic patterns of year level means require more detailed investigation to determine why these patterns exist. It would appear that longitudinal studies are required in order to do this thoroughly.

DISCUSSION

The results of this major trial of the LPSS with Australian school children would appear to indicate quite clearly that expressed student preferences for co-operative, competitive, and individualised learning modes are independent. Thus, a student may be positively disposed to working in any, all, or none of these learning modes. Furthermore, the content, factorial, and correlational validity data obtained in this study support the use of the LPSS (Form B) in school settings. The internal consistency and stability coefficients would also appear to be adequate to justify the

FIGURE 1

STANDARDISED MEAN SCORES ON LPSS Co-OPERATIVE, COMPETITIVE, AND INDIVIDUALISED SUB-SCALES FOR BOYS AND GIRLS WITHIN YEAR LEVELS

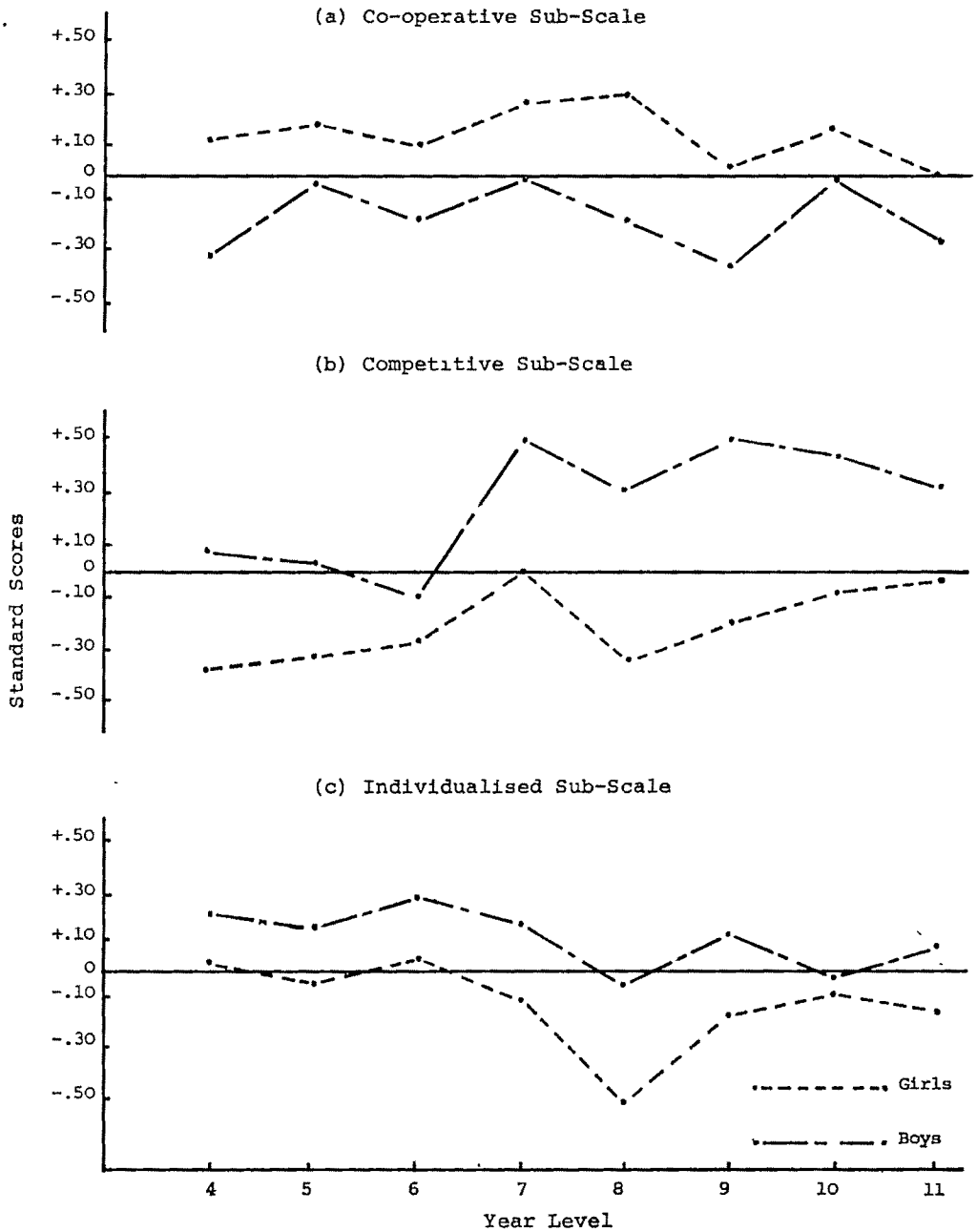


TABLE 6

SCHOOL YEAR LEVEL $\times$ SEX ANALYSES OF VARIANCE OF THE CO-OPERATIVE, COMPETITIVE, AND INDIVIDUALISED SUB-SCALE SCORES (MAIN TESTING)

Source of Variance	df	Co-operation		Competition		Individualised	
		MS	F	MS	F	MS	F
Year level	7	83	2.3*	253	6.5**	172	4.1**
School year	1	1463	40.0**	3500	90.0**	1039	24.9**
Sex	7	25	0.7	63	1.6	25	0.6
Year $\times$ Sex	1625	37	—	.39	—	42	—
Residual							

* $P < 0.05$.** $P < 0.001$.

use of the LPSS in school decision-making. However, it is desirable and now possible, using the results of the present study, to develop a shortened version of the LPSS which might be used more easily in schools and work on such an instrument has been undertaken.

Following the extensive aptitude-treatment interaction research of the last decade (Cronbach and Snow, 1977), it has now been suggested that the use of temperament-treatment interactions may allow for greater efficiency in learning processes (Barclay, 1978). The basic point is that there may be a single teaching style that is best suited to the cluster of attributes that forms an individual student's personality, or alternatively, that there is no single teaching style that will work to good advantage with all students' temperaments. The problem is how to make the appropriate match between teaching style and student temperament and do it practically for large numbers of students in institutions with complicated timetabling problems for the scheduling of classes.

The difficulty is considerable, and yet the problem is not beyond solution. In one American setting it has been approached on a very ambitious scale. With the aid of substantial technological support, it has been found possible at least to provide teachers with a dossier of group and individual data on the cognitive, social, and attitudinal attributes of both individual students and classroom groups (Barclay, 1977a, 1977b). By means of computer-composed reports based upon the Barclay Classroom Climate Inventory (BCCI), individual students are described and classroom groups are identified in terms of their likely susceptibility to a particular teaching style. Obviously on a much smaller scale, the LPSS can be seen as an aid to such an end.

If a teacher has a means of discovering the learning preferences of the students in the class, then certain curriculum decisions may be much more easily made, e.g.:

Which students might productively be grouped for project work?

Which students might be assigned to individual library research?

Which students might strive to excel whenever given the opportunity?

Might two learning modes be set up and monitored in the classroom at the same time, e.g., co-operative and individualised?

Are the expressed preferences true for a range of school subjects or for only a few, perhaps not including the one in question?

How much time might be given to competitive learning games before students' interest and involvement fall away?

The aim in any particular case would be to recognise the existence of alternate styles and differing student preferences and to attempt a match between these. The crucial

point is that the use of the Barclay BCCI in all its complexity, or the LPSS in its relative simplicity, may place in the hands of teachers some information on which to base a 'local instructional theory' (Snow, 1977b). Such a theory, it is argued, being locale-specific and easily modified by formative evaluation, has a greater prospect of success with *these* students, in *this* place, at *this* time than a general instructional theory intended to function satisfactorily in many instances and settings, and with a variety of teaching practitioners. The information gathered for or by teachers can be both aptitude and temperament data. The evidence is that aptitude-treatment interactions in learning exist, and that they matter (Snow, 1977a); the evidence is only beginning to be catalogued for temperament-treatment interactions (Barclay, 1978), but this task needs to be pursued without delay.

ACKNOWLEDGMENT.—This paper is produced as part of the Co-operative, Competitive, and Individualised Learning in Education Project (CCILE). The CCILE Project has been funded by the Education Research and Development Committee, Canberra, and has received the full co-operation of the New South Wales Department of Education. The authors gratefully acknowledge the research assistance of Jennifer Barnes in this project.

REFERENCES

- BARCLAY, J. R. (1977a). Letters from a computer friend: a novel way to help kids. Unpublished mimeographed report, University of Kentucky.
- BARCLAY, J. R. (1977b). *Providing for Individual Differences: a Manual of the Barclay Classroom Climate Inventory* (4th edition). Lexington, Kentucky: Educational Skills Development.
- BARCLAY, J. R. (1978). Temperament clusters and individual differences in the elementary classroom: a summary. Paper presented at the Annual Conference of the American Educational Research Association, Toronto.
- BARNES, J., OWENS, L., and STRATON, R. (1977). *Learning Preference Scale—Students, Form B*. University of Sydney, Department of Education.
- CONNELL, W. F., STROOBANT, R. E., SINCLAIR, K. E., CONNELL, R. W., and ROGERS, K. W. (1975). *12 to 20: Studies of City Youth*. Sydney: Hicks Smith.
- CRONBACH, L. J., and SNOW, R. E. (1977). *Aptitudes and Instructional Methods: A Handbook for Research on Interactions*. New York: Irvington.
- DAVIS, J. R., and SPEARRITT, P. (1974). *Sydney at the Census, 1971: A Social Atlas*. Canberra: Urban Research Unit, Australian National University.
- DEUTSCH, M. (1949a). A theory of cooperation and competition. *Human Relations*, 2, 129-152.
- DEUTSCH, M. (1949b). An experimental study of the effects of cooperation and competition upon group process. *Human Relations*, 2, 199-232.
- JACKSON, D. H. (1971). The dynamics of structured personality tests: 1971. *Psychol. Rev.*, 78, 229-248.
- JOHNSON, D. W. (1974). Evaluating affective outcomes of schools. In WALBERG, H. (Ed.), *Evaluating School Performance*. Berkeley, California: McCutchan.
- JOHNSON, D. W., and AHLGREN, A. (1976). Relationship between student attitudes about co-operation and competition and attitudes toward schooling. *J. educ. Psychol.*, 68, 92-102.
- JOHNSON, D. W., and JOHNSON, R. T. (1974). Instructional goal structures: co-operative, competitive, or individualistic. *Rev. educ. Res.*, 44, 213-240.
- JOHNSON, D. W., and JOHNSON, R. T. (1975). *Learning Together and Alone: Co-operation, Competition, and Individualization*. Englewood Cliffs: Prentice-Hall.
- JOHNSON, D. W., and JOHNSON, R. T. (1978). The instructional use of co-operative, competitive, and individualistic goal structure. In WALBERG, H. (Ed.), *Educational Environments and Effects*. Berkeley, California: McCutchan.
- JOHNSON, D. W., and JOHNSON, R. T. (1979). Conflict in the classroom: controversy and learning. *Rev. educ. Res.*, 49, 51-69.
- JOHNSON, D. W., JOHNSON, R. T., and ANDERSON, D. (1978). Relationship between student co-operative, competitive, and individualistic attitudes and attitudes toward schooling. *J. Psychol.*, 100, 183-199.
- JOHNSON, D. W., and NOREM-HEBEISEN, A. A. (1977). Attitudes toward interdependence among persons and psychological health. *Psychol. Rep.*, 40, 843-850.
- JOHNSON, D. W., and NOREM-HEBEISEN, A. A. (1979). A measure of co-operative, competitive, and individualistic attitudes. *J. soc. Psychol.*, 109, 253-261.
- NIE, N. H., HULL, C. H., JENKINS, J. G., STEINBRENNER, K., and BENT, D. H. (1975). *SPSS: Statistical Package for the Social Sciences*. New York: McGraw-Hill.
- NOREM-HEBEISEN, A. A., and JOHNSON, D. W. (1977). The relationship between co-operative,

- competitive and individualistic attitudes and differentiated aspects of self-esteem. Unpublished mimeographed report, University of Minnesota.
- SNOW, R. E. (1977a). Research on aptitude for learning: a progress report. In SHULMAN, S. (Ed.). *Review of Research in Education*, 4 (1976). Itasca, Illinois: Peacock.
- SNOW, R. E. (1977b). Individual differences and instructional theory. *Educ. Researcher*, 6 (10), 11-15.

(Manuscript received 26th October, 1979)

ACHIEVEMENT PERFORMANCE OF INTERNALLY AND EXTERNALLY ORIENTED BLACK AND WHITE HIGH SCHOOL STUDENTS UNDER CONDITIONS OF COMPETITION AND CO-OPERATION EXPECTANCIES*

By P. S. FRY

(The University of Calgary)

AND K. J. COE

(Calgary General Hospital)

SUMMARY. In order to test the hypothesis that factors of racial differences may mediate the locus of control-achievement relationship, high school students were asked to co-operate with or compete against partners of within-race or opposite race. As hypothesised, the achievement behaviour of internally oriented black students was more influenced by the race of the interactor and the competitive-co-operative expectancies in the task than that of the white students. While the achievement performance of internally oriented blacks improved more than did the achievement of external blacks, internal blacks increased their performance most when competing against whites or when co-operating with other blacks. The findings have implications for classroom practice and achievement performance, especially in schools that have desegregated multiracial programmes.

INTRODUCTION

As Pettigrew (1967) observes, recent socio-economic and academic advances are being viewed by blacks "as gains due to their increasing sense of internal control over their lives and not simply as gains granted coincidentally or paternalistically by outside agencies". Forward and Williams (1970) view the concept of internal and external control as being the most central in understanding the psychological bases for the reaction of young negroes towards the racist attitudes and practices of the dominant whites. Presenting data from high school and college samples Gurin *et al.* (1969) showed that black students endorsed the internal control ideology equally as strongly as white students and were effectively moving away from an external orientation or an individual blame system to an orientation of personal efficacy. To relate the internal-external concept more specifically to the racial situation, the Gurin *et al.* (1969) data showed that a high score on the internal orientation was associated with higher achievement, greater participation and the belief that co-operative action among blacks is needed to combat white dominance and racial prejudice. As Weigel *et al.* (1975) observe, in order to maximise the beneficial outcomes of desegregation in schools and to minimise the negative correlates, it is essential that we recognise a number of specifically social-psychological concepts such as an increasing sense of self-esteem, social comparison and internal-external control that may mediate the reactions and responses of the young black in cross-race and cross-ethnic interactions. One important condition for constructive race relations cited by reviewers (e.g., Gerard and Miller, 1975) is the creation of interdependencies among students from various racial groups. In this regard, co-operative and competitive factors have long been thought to be important aspects of cross-ethnic contacts (Allport, 1954; Harding *et al.*, 1969; Cohen, 1973). Gottlieb (1965) and Cusick (1973), for example, found considerable within-race preference for task relationships and classmates by both black and white. Following the hypothesis of the Blocked Opportunity Theory postulated

* Data for this research were collected while the first author was on sabbatical leave at the Pennsylvania State University.

by Caplan and Paige (1968) and Caplan (1970), several reviewers agree that negroes who have high aspirations for their own lives and believe in their ability to achieve these goals are less likely to co-operate with their white counterparts in cross-ethnic intergroup dependencies. A selective interpretation of Ransford's (1968) theory of Alienation-Powerlessness also provides support for the prediction that the negroes having lived in structural isolationism and acquiescence for a prolonged time are now more likely to use unconventional means of non-co-operation to express grievances against their white counterparts.

The present study was therefore undertaken to examine systematically the differential effects of co-operation and competition expectancies on the achievement of internally and externally oriented black and white students. While there is some clear evidence for a relation between an internal locus of control orientation and greater academic achievement for whites (Allen *et al.*, 1974; Prociuk and Breen, 1974; Parent *et al.*, 1975; Daniels and Stevens, 1976) there are little data available on the consistency of a relationship between locus of control and achievement for blacks. In the present study the authors examined the possibility that the locus of control/achievement relation may be mediated by certain racial preferences. Two prime variables that were expected to be mediators of the locus of control/achievement relation were race of the interactors and the co-operative *versus* competitive nature of the achievement task. There is some reason to believe that achievement performance of some blacks may suffer when they compete with whites (Amir, 1969; Pettigrew, 1969; Cohen, 1973; Weigel *et al.*, 1975). Likewise, some work (e.g., Staub, 1971; Johnson and Johnson, 1974) generally suggests that whites are more responsive to competition, and by contrast blacks are more responsive to co-operation in terms of achievement performance (Thelen, 1970; Johnson and Johnson, 1972; Watson and Johnson, 1972).

Thus, in order to investigate the effects of the degree of competition with the race of the subjects and their interactive partners on achievement behaviour, male and female adolescents differing in locus of control orientation competed against or co-operated with same-or opposite race partners. Rotter's social learning theory (Rotter *et al.*, 1972) was the source for the concept of locus of control as employed in the present study. It was predicted that while internals generally would show more achievement effort than externals, we also predicted that internal blacks would be more sensitive to changes in environmental expectancies than the internal whites. Overall, the prediction was that the responses of internal blacks to expectancies of co-operation and competition would be different from those of internal whites, since the internal orientation of the former is not consistent with the traditional black role of passivity and co-operativeness. In other words, the expectation was that the achievement of internal blacks would be modified by whether they were competing or co-operating with same-race or opposite-race partners. Based upon the recent tendency of blacks to show more incentive to achieve against whites it was hypothesised that internal blacks would probably increase their achievement behaviour most when competing with whites than with blacks.

METHOD

Sample

The sample consisted of 80 black (20 females and 60 males) and 80 white (30 females and 50 males) students (median age = 17.8 and 17.4 years respectively) all drawn from a much larger pool of students attending two major high schools in Pittsburgh, and another high school in a suburban community of Pittsburgh. Both communities were primarily lower class to lower middle class in socio-economic level. The schools had approximately 40 per cent black, 50 per cent white and 10 per cent Hispanic populations. Since our measures were not relevant to a Hispanic population, Hispanic surnamed children were eliminated from the study. A large percentage of the

black and white children who obtained parental approval to participate in the study were accepted for the experiment.

Materials

Locus of control orientation was measured by an adapted version of the Gurin *et al.* (1969) Locus of Control scale in which some of the items were simplified so as to be comprehensible to a high school population of students. The Gurin *et al.* adaptation of the Locus of Control scale was selected for use in this study because several items of self-other orientation, personal control and control ideology were written to measure beliefs about the role of external and personal forces in the race situation. However, items which make a direct reference to the negro race had to be eliminated. Included were 23 items that Gurin *et al.* borrowed from the Rotter Internal-External Control scale (Rotter, 1966), three items from Gurin *et al.*'s (1969) Personal Efficacy scale and six items that focused on the respondent's race ideology. The six filler items in Rotter's (1966) scale were also retained. Thus 32 questions about internal and external control were posed in a manner that was meaningful to the control ideology and personal control of both whites and negroes.

The main performance task chosen for the present study was a digit symbol task adapted from the Wechsler Adult Intelligence Scale (1955). The test has a key in which digits are associated with separate symbols. The subject's task was to match appropriate symbols with the digits in a randomly ordered presentation of 100 digits. The subject had 120 seconds to complete as many of the symbols as possible. While this task was not strictly comparable with the criterion of school achievement, it was selected because it is a non-verbal test which is relatively subculture free and is seldom used to screen students in school admission or job placement. Experimenters were a white male graduate student and a black male graduate student. Subjects were assigned to them in random order.

Procedure

In the first session, subjects were administered the Locus of Control measure and the digit symbol task in groups of 15 to 20 students. The purpose of the tests as presented to the subjects was to test the relationship between subjects' attitudes and their performance in the digit task. Subjects were informed that they would be called back for a second administration of the tests and their co-operation was solicited. Complete randomisation of the order in which individual subjects took the two tests was not possible. However, half of the subjects took the digit task first and the other half took the Locus of Control measure first.

On the basis of the Locus of Control scores, all black and white subjects were divided into internal and external groups. The median score of 16 was used as a cut-off point with scores less than 16 being considered internal and scores of 16 or more being considered external.

After about two weeks, individual subjects were asked to come back for a second test. Subjects in the internal and external groups were randomly assigned to a competitive or co-operative condition. In the experimental conditions the procedure was designed to maximise the degree of competitiveness and co-operativeness through situational stress and by varying the instructional set.

In all conditions, subjects were told that their partners were subjects like themselves who had participated in the initial testing. In the competitive condition, the subject was led to believe that he/she and the partner got an identical score on the digit test. The subject and the partner were then told: "I would like you to take the test again so that we can break the tie and get a separate rank for each of you. It is very important that you do your best to excel, since whoever gets the highest score on this test will receive a reward of ten dollars and extra credit in school work." By

contrast, subjects in the co-operative condition were told that in the initial testing the scores were based on individual performance, and in the retest, the subject's score would be added to his/her partner's score in arriving at a cumulative score. The subjects were again encouraged to do their very best "since the pair that receives the highest combined score will receive a reward of ten dollars to be equally split between the two, and extra credit in school work".

The digit-symbol test was readministered with a time limit of 120 seconds, as in the original testing. In order to allow a larger number of subjects to be tested, black and white male and female confederates of the experimenter acted as the subject's partner in half of the cases. The confederates were two black and two white (one male, one female in each race) students who were undergraduates at the university and who were sufficiently young-looking to pass off as being equal in age to the subjects. None of the subjects had met the experimenter's accomplices before. These accomplices received prior training in enacting the co-operative and competitive role.

RESULTS

For each subject, the difference between the scores on the first and second digit symbols test was obtained. The mean difference scores are shown in Table 1. The presentation of the experimental stimuli did not show an order effect, nor were there

TABLE 1

MEAN DIFFERENCES BETWEEN FIRST AND SECOND ADMINISTRATION OF THE ACHIEVEMENT TASK FOR INTERNALLY AND EXTERNALLY ORIENTED BLACK AND WHITE SUBJECTS IN CO-OPERATIVE AND COMPETITIVE CONDITIONS

Partner	Internal		External	
	Black	White	Black	White
<i>Same Race</i>				
Co-operative condition	10.61	9.11	Co-operative condition	6.57
Competitive condition	6.24	8.44	Competitive condition	5.95
<i>Different Race</i>				
Co-operative condition	3.45	8.99	Co-operative condition	1.91
Competitive condition	11.95	7.45	Competitive condition	5.35

significant effects as a function of sex of subjects' or experimenter's accomplices. There was no main effect of race of experimenter ($F < 1$). As a result, the data were collapsed across these variables for analysis. The results were analysed by a 2(Internal vs. External) $\times$ 2(Black vs. White subject) $\times$ 2(Black vs. White partner) $\times$ 2(Co-operation vs. Competition) analysis of variance.

Consistent with the basic hypothesis, internals increased their performance significantly more than did externals, $F(1, 143) = 10.31$, $P < 0.01$. However, there was a significant three-way interaction of Internal vs. External $\times$ Race of Partner $\times$ Co-operation vs. Competition, $F(1, 143) = 12.51$, $P < 0.01$. Newman-Keuls procedures indicated that the interaction came from internal blacks co-operating with a black or competing with a white, thus increasing their performance significantly more than internals co-operating with whites or competing with blacks, or external blacks co-operating with whites.

DISCUSSION

The investigation of the locus of control/achievement relationship of black students is predicated on the assumption that their achievement behaviour emerges from a number of complex interactions (Epps, 1969). The importance of the internal-

external concept for understanding the achievement of blacks was confirmed. However, the results suggest that the locus of control/achievement relationship for blacks is not a completely straightforward one as exists for the white internals and externals. While internal blacks are more achieving than external blacks yet even the internal blacks are not consistent in their achievement and show different levels of performance depending upon whether they are co-operating with blacks or competing against whites.

The findings of the present study confirmed that the important variables which served as mediators of the locus of control/achievement relationship for the internal black student were within-race or opposite race characteristics of the interactors and the competitive or co-operative format of the achievement task. The internal blacks performed best when competing against whites or co-operating with blacks against other interactors. In achievement situations such considerations underscore the possibility that putting the internal black in competition with a black would be counter-productive but putting an internal black in competition with a white would induce motivation. The starting point, of course, would be a knowledge of the locus of control orientation of the interactors which may suggest the most productive procedures for matching black and white students in learning situations. Although a direct application of this procedure to a public school setting is not possible, it should be feasible to develop an analogous approach in which the internal black competes with the internal white and the internal black is paired with another black in those situations of learning and performance in which a co-operative, supportive relationship might be conducive to improving the performance of an external black, or where two internal blacks can improve team effort. In interpreting the results of this analogue study it must be kept in mind that the black and white partners' exposure to each other was very brief. This variable limits the generalising of the findings to naturalistic settings until replications are conducted with black and white subjects having more prolonged interactions with one another.

Aside from the comparative outcomes of the internal and external control strategies, our results raise some provocative questions about the conditions under which the black is willing to demonstrate his/her true ability. There appears to be an interesting race-role reversal in which the black's traditional rejection of his/her own race and identification with the white is no longer apparent (cf. discussion in Katz, 1969), and the identity conflict is slowly phasing out. The results of the present study suggest that the internal blacks, perhaps because of a growing sense of personal efficacy and competence, are co-operative with rather than rejecting of other blacks. More research is needed to assess whether these findings are specific to high-school age blacks or whether they represent an on-going trend in cross-ethnic interactions.

[Requests for reprints should be sent to P. S. Fry, The University of Calgary, Department of Educational Psychology, Calgary, Alberta, Canada, T2N 1N4.]

REFERENCES

- ALLEN, G. J., GIAT, L., and CHERNEY, R. J. (1974). Locus of control, test anxiety, and student performance in personalized instruction course. *J. educ. Psychol.*, 66, 968-973.
- ALLPORT, G. (1954). *The Nature of Prejudice*. Cambridge, Mass.: Addison-Wesley.
- AMIR, Y. (1969). Contact hypothesis of ethnic relations. *Psychol. Bull.*, 71, 319-343.
- CAPLAN, N. (1970). The new ghetto man: a review of recent empirical studies. *J. social Issues* 26, 59-73.
- CAPLAN, N. S., and PAIGE, J. M. (1968). A study of ghetto rioters. *Scient. Am.*, Aug., 219 (2), 15-21.
- COHEN, E. G. (1973). Modifying the effects of social structure. *Am. behav. Scient.*, 16, 861-879.
- CUSICK, P. A. (1973). Black-white student interactions in an urban secondary school. Paper presented at the annual convention of the American Educational Research Association, New Orleans, February.

- DANIELS, R. L., and STEVENS, V. P. (1976). The interaction between the internal-external locus of control and two methods of college instruction. *Am. educ. Res. J.*, 13, 103-113.
- EPPS, E. G. (1969). Negro academic motivation and performance: an overview. *J. social Issues*, 25, 5-11.
- FORWARD, J. R., and WILLIAMS, J. R. (1970). Internal-external control and black militancy. *J. social Issues*, 26, 75-92.
- GERARD, H. B., and MILLER, N. (1975). *School Desegregation: A Long Range Study*. New York: Plenum Press.
- GOTTLIEB, D. (1965). School integration and absorption of newcomers. *Integrated Education*, 3, 69-75.
- GURIN, P., GURIN, R., LAO, R. C., and BEATTIE, M. (1969). Internal-external control in the motivational dynamics of Negro youth. *J. social Issues*, 25, 29-54.
- HARDING, J., PROSHANSKY, H., KUTNER, B., and CHEIN, I. (1969). Prejudice and ethnic relations. In LINDZEY, G. and ARONSON, E. (Eds.), *Handbook of Social Psychology*, (Vol. 5). Reading, Mass.: Addison Wesley.
- KATZ, I. (1969). A critique of personality approaches to Negro performance, with research suggestions. *J. social Issues*, 25, 13-27.
- JOHNSON, D. W., and JOHNSON, R. T. (1974). Instructional goal structure: co-operative, competitive, or individualistic. *Rev. educ. Res.*, 44, 213-240.
- JOHNSON, S., and JOHNSON, D. W. (1972). The effects of others' actions, attitude similarity and race on attraction towards the other. *Hum. Relat.*, 25, 121-130.
- PARENT, J., FORWARD, J., CENTER, R., and MOHLING, J. (1975). Interactive effects of teaching strategy and personal locus of control on student performance and satisfaction. *J. educ. Psychol.*, 67, 764-769.
- PETTIGREW, T. F. (1967). Social evaluation theory. Convergences and applications. In LEVIN, D., (Ed.), *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press.
- PETTIGREW, T. F. (1969). The negro and education: problems and proposals. In KATZ, I., and GURIN, P. (Eds.), *Race and the Social Sciences*. New York: Basic Books.
- PROCIUK, T. J., and BREEN, L. J. (1974). Locus of control, study habits and attitudes and college academic performance. *J. Psychol.*, 88, 91-95.
- RANSFORD, H. E. (1968). Isolation, powerlessness and violence. A study of attitudes and participation in the Watts riot. *Am. J. Sociol.*, 73, 581-591.
- ROTTER, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychol. Monogr.* 80 (1, Whole No. 609).
- ROTTER, J. B., CHANCE, J. E., and PHARES, J. E. (1972). *Applications of a Social Learning Theory of Personality*. New York: Holt, Rinehart and Winston.
- STAUB, E. (1971). Helping a person in distress. The influence of implicit and explicit "rules" of conduct on children and adults. *J. Person. soc. Psychol.*, 17, 137-144.
- THELEN, H. A. (1970). A proposal for the attainment of racial integration through public education. *School Rev.*, 391-396.
- WATSON, G., and JOHNSON, D. W. (1972). *Social Psychology: Issues and Insights* (2nd ed.). Philadelphia: Lippincourt.
- WECHSLER, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.
- WEIGEL, R. H., WISER, P. L., and COOK, S. W. (1975). Impact of co-operative learning experiences on cross-ethnic relations and attitudes. *J. social Issues*, 31, 219-245.

(Manuscript received 21st March, 1979)

USING STUDENT RATINGS AND CONSULTATION TO IMPROVE INSTRUCTION

By W. J. McKEACHIE, Y-G. LIN, MONICA DAUGHERTY,
MARY M. MOFFETT, CYNTHIA NEIGLER, J. NORK, MARTINA WALZ,
AND R. BALDWIN

*(Department of Psychology, and Centre for Research on Learning and Teaching,
University of Michigan)*

SUMMARY. The effect upon teaching of feedback from student ratings was compared with the effect produced when such feedback was supplemented by consultation from a more experienced teacher. Consultation enhanced the positive effect of the feedback in terms of later student ratings. The effect upon student achievement in the classes of the teachers involved tended to be positive but not consistent across courses.

INTRODUCTION

In a study of the effects of feedback of pupil ratings on elementary school teachers' behaviour, Gage *et al.* (1963) found that sixth grade teachers changed their behaviour in the direction of pupils' descriptions of the ideal teacher. Tuckman and Oliver (1968) found at the high school level that feedback from student ratings changed teachers' behaviour significantly as compared with change in the behaviour of teachers not receiving feedback. Marsh *et al.* (1975) also found that feedback from student ratings resulted in improved ratings at the end of the semester. Miller (1971), however, found no significant overall effect of feedback of mid-semester college student ratings. In one of the three courses involved, the feedback was associated with significantly higher student achievement on the final examination. Miller concluded that "... the results suggest some limitations in the use of student ratings as a method of improving instruction". Centra's results (1973) support this conclusion.

A previous finding from our own research programme helps explain the previous results and points toward the need for further research. In this study Pambookian (1974) also found little overall effect of feedback of student ratings, but detailed analysis of his data indicated that significant positive changes did occur for those instructors in the mid-third of the distribution of initial ratings. Pambookian suggests that those instructors in the top third had little need to improve while those in the bottom third may have become more anxious and defensive as a result of feedback.

What can be done to increase the usefulness of student ratings? We theorised that the effectiveness of feedback depends upon at least three variables:

- (1) Does the feedback convey information the subject does not know?
- (2) Does the feedback increase or decrease motivation?
- (3) Does the performer know what to do to improve?

We believed that the effectiveness of feedback dependent upon the last two of these variables could be helped by personal consultation, a belief supported by Aleamoni's (1974) results.

METHOD

Hypothesis

Our hypothesis was that teachers given feedback of student ratings of instruction with encouragement and suggestions for improvement would be superior on end-of-term measures of teaching effectiveness to teachers receiving printed feedback of student ratings or those receiving no feedback.

Sample and research design

The sample consisted of 37 graduate student teaching fellows and three faculty members recruited from among the individuals teaching in the introductory psychology courses at the University of Michigan. The 40 teachers were assigned to one of three groups which received:

- (1) personal feedback of student ratings with consultation from an experienced teacher, or
- (2) computer print-out of student ratings, or
- (3) no student ratings.

Teachers were assigned randomly within each course and within three levels of teaching experience:

- (1) those with no experience,
- (2) experienced teaching fellows, and
- (3) professors.

Thus one full professor was assigned to each kind of group.

Of the 37 teaching fellows, 21 had had at least one year of previous experience; all three faculty members were full professors with 15 to 40 years of experience. All participants were rated favourably by their students in both the pre-test and post-test, and pre-test group means were equal. Twenty-two of the teachers were men; 18 were women.

Teachers were informed that we were interested in studying the effect of student ratings on teaching and that we would provide feedback from student ratings to those in the two experimental groups.

Class sizes ranged from 10 to 35 and mean class size did not differ among the three conditions. The proportion of students filling out the questionnaires in each class averaged 80 per cent, with no significant differences between groups in rate of response.

Measures

(a) *Student perception of teaching and learning.* The primary measure was the Michigan Student Perception of Teaching and Learning form. This form has evolved over a period of 20 years in which items from the major forms used in the 1950s for student ratings of teaching were factor analysed to obtain items best representing the major dimensions used by students in rating teachers (Isaacson *et al.*, 1964). The form was then revised as the result of a series of validity studies (McKeachie *et al.*, 1971) and of a multiple discriminant analysis (McKeachie and Lin, 1975). The version of the form used in this study consisted of 32 items. The form administered as a post-test differed from the pre-test form only in containing two additional items requiring evaluation of the instructor's general teaching effectiveness and the value of the course as a whole.

Three scores from this form were used as dependent variables:

- (1) Rating on the item, "Rate the instructor's general teaching effectiveness for you."
- (2) Rating on the item, "Rate the value of the course as a whole to you."
- (3) Sum of scores on items with high loadings on the factor, "Impact on Students."

(b) *Psychological thinking.* While most other studies of feedback of student ratings have dealt with the effect of feedback on later student ratings, the ultimate criterion of teaching effectiveness is not student ratings but student learning. An important outcome measure, therefore, was student performance on items selected from the Introductory Psychology Criteria Test (Milholland, 1964). This test was developed by a committee of the Division on Teaching of the American Psychological

Association and has been used in much previous research in introductory psychology courses. It provides a measure of achievement of the higher level cognitive outcomes of the Bloom Taxonomy. Teachers involved in the research were not familiar with the test and did not teach to the test. It thus represents a measure unbiased with respect to the specific content emphasised by different teachers.

(c) *Attitudes.* An 'Attitude toward Psychology' Questionnaire consisting of eight Likert-type items drawn from Carrier's scale (1966) was administered as part of the test battery. Since previous use of the test had indicated that some of the test items appeared to involve an uncritical, strong endorsement of psychology, the scoring key was changed to indicate simply agreement, disagreement, or undecided. Only those items were used which five introductory psychology teachers agreed to be appropriate in terms of the goals of introductory psychology courses. A locally constructed 10-item Likert type scale of Attitude toward Self was used to assess impact of the introductory classes on the student as a person.

Procedures for feedback

Members of the research staff administered the Michigan Student Perception of Teaching and Learning from during the last 15 to 20 minutes of a class period during the fifth to seventh weeks of the 14-week term. Wording of some items was changed to the present tense from the past tense to make the form appropriate for administration early in the course rather than at the end of the course.

Rating forms were scored by computer. The computer print-out presented the mean of the class and the mean of all classes in the same course for each of the seven dimensions for which the form is scored: Impact on Student, Rapport, Standards, Group Interaction, Difficulty, Structure, and Feedback. Individual class and course means for each item were presented following the factor scores.

The print-outs were mailed during the fifth to eighth weeks of the term to those in the 'Printed Feedback' group and during the same period they were reviewed personally with each teaching fellow in the group receiving personal feedback.

At the beginning of the consultation sessions teachers were asked to fill out forms indicating their expectation of the student ratings on each dimension, their own self-perceptions and where they would like to be. Typically, the consultant then asked them how the class was going and, in response to their reactions, suggested how the student ratings confirmed (or rarely, did not confirm) their perceptions. He then pointed out factors on which the teacher differed significantly from the mean of all classes. If there seemed to be any problems, he suggested some possible alternative methods of handling the problem. As indicated earlier, the mean ratings in all groups were relatively favourable, so that the hope that the consultant could help teachers cope with very negative feedback was not realised.

Post-test

During the eleventh and twelfth weeks of the 14-week term students were invited by postcard to attend evening sessions to take the tests measuring the dependent variables, including the post-test of student perception of learning and teaching. On the average the number of students per class taking the post-test was one less than the number of students taking the pre-test, and did not differ significantly between groups. Upon completion of the tests each student received written feedback describing the experimental design and the nature of the tests used.

RESULTS

Two subsidiary analyses were also carried out. Since all classes were relatively small, we had anticipated no effect of class size and found none. We also found no significant difference in pre-test ratings between those groups pre-tested.

We also asked whether those teachers who changed the most were those most effective in terms of student performance on our criterion measure. Correlating positive change on the seven dimensions of the student rating form with mean student scores on our four criterion measures, we found that 26 of the 28 correlations were positive, with five of the seven correlations with the Attitude toward Self measure being significant at the 0.05 level or better. Thus those teachers who were perceived as improving were also most effective in terms of student performance.

Experimental studies of classes and teachers almost invariably encounter the problem of a small number. While we have chosen to use analysis of variance, we suspect that visual inspection of the cell means will be more informative to the reader than the F tests.

As Table 1 indicates, the student ratings at the end of the term support the primary hypothesis. Those instructors who received consultation about their student ratings earlier in the term were rated as most effective at the end of the term.

TABLE 1
EFFECTS OF FEEDBACK ON TEACHING EFFECTIVENESS: STUDENT PERCEPTIONS OF TEACHERS

	Mean class scores			F	P
	Consultation with student ratings	Printed feedback of ratings	No feedback		
<i>Overall items*</i>					
General teaching effectiveness	3.6	3.1	3.0	4.55	0.017
Overall value of course	3.4	3.0	3.0	5.77	0.007
Impact on students†	4.17	3.95	3.80	3.60	0.037
Number of sections	14	13	13		

* Scale for overall items: 5 = Excellent; 4 = Very Good; 3 = Good; 2 = Fair; 1 = Poor.

† A high factor score represents high impact as perceived by students.

The effect of consultation upon effectiveness as measured by student achievement was not as clear-cut. As Table 2 indicates, the hypothesis received statistically significant support in terms of student attitude toward the self. On the other criteria the combined results were not significantly different. In one of the four courses, however, students in the classes of teachers receiving counselling were significantly superior on the measure of psychological thinking. Although the significant differences found were in the hypothesised direction, they were not consistent across courses.

TABLE 2
EFFECT OF FEEDBACK ON TEACHING EFFECTIVENESS: MEASURES OF COURSE GOALS

Variables	Consultation with student ratings	Printed feedback of ratings	No feedback	F	P
Psychological thinking	50.29*	49.71	49.98	1.20	0.31
Attitude toward psychology	10.14	10.32	9.77	0.78	0.46
Attitude toward self	4.88	4.39	4.25	3.16	0.054
N	14	13	13		

* Standardised scores. All scores in this table are mean scores across courses. On the measure of psychological thinking one course had a significant difference favouring the group with consultation.

Which teachers are helped by consultation?

We had expected the effects of consultation to be most helpful to those teachers with poor ratings. Pambookian (1974) had shown printed feedback to be helpful to those in the middle third of his distribution but suggested that those in the lower third of the distribution might be discouraged by feedback. Our personal consultation, we thought, would help reduce the potential negative effects of negative ratings by reducing anxiety, by increasing motivation to improve through increasing hope of success, and by suggesting alternative behaviours to those criticised. We thus separated the groups into thirds for further analyses. As Table 3 indicates, the results were in the direction predicted, but the differences were small. A similar analysis of the other criterion measures did not reveal any consistent patterns.

TABLE 3
POST-COURSE STUDENT RATINGS OF IMPACT FOR TEACHERS DIFFERING ON INITIAL RATINGS

Initial ratings	Consultation with student ratings			Printed feedback of ratings		
	Mean	SD	N	Mean	SD	N
Low	3.9	0.21	5	3.5	0.44	4
Mid	4.2	0.19	4	4.0	0.08	4
High	4.5	0.27	5	4.3	0.31	5

$t = 1.02$ for differences of differences. Low vs. mid and high.

DISCUSSION

This study supports our original hypothesis. Simple knowledge of student ratings is not as effective in enabling a teacher to improve as is the same information presented with encouragement and suggestions. (In fact, in a later experiment we again found that the major part of the improvement was due to consultation rather than to the student ratings.) As a result of our experience in these two experiments, it seems to us that consultation provides motivational support as well as opportunities for rethinking the teaching situation. Further research should be directed toward studying what factors influence motivation for improvement following feedback. Since most teachers are rated favourably, motivation derived from a sense of basic competence and desire to enhance that competence is likely. But what of those who receive negative feedback? Motivation to avoid teaching, anxiety about classroom interactions, concern about loss of status—these, and other motives unlikely to contribute to improved teaching, are likely to ensue. Thus the task of the consultant may be quite different in the two situations.

Sullivan (1979) has pointed out that the process of improvement is different for individuals at different levels of teaching ability. For those who are poor teachers, improvement involves diagnosis and elimination of serious weaknesses or errors, suggestion of alternative approaches and encouragement. For those who are good or excellent teachers, improvement may involve developing clearer conceptualisations of teaching, learning of student characteristics, increasing the teacher's repertoire of skills, providing better tools, or refining existing skills. Since the ratings indicate that our teachers were in this latter group, the diagnostic potential of ratings was not so important as it might be in counselling less effective teachers.

The problem of enhancing the effectiveness of student ratings for improving teaching is related to the broader theoretical problem of how people in general use statistical information. Our studies of the impact of student ratings upon teaching support those of Borgida and Nisbett (1977) on the impact of abstract vs. concrete information upon decisions. Borgida and Nisbett found that consensus information

is not as effective an influence on attribution processes or behaviour as would be expected on the basis of theory. They reject Kahneman and Tversky's (1973) suggestion that people lack cognitive structures for dealing with statistical information. Rather, they say, the problem is that the information is too pallid and abstract. People generally are not moved by statistical information, whether it be about the mortality rates of smokers or about the quality of teaching in a college course.

Studies of the impact of statistical information are of major significance for those concerned with student ratings of teaching. Most of the research on student ratings has been directed toward establishing for what ratings are valid and how the quality of the information can be improved—as, for example, through the use of factor analyses to identify the minimum number of items conveying the maximum information about the major dimensions of teaching.

The Michigan researchers have carried out several studies of the impact of student ratings upon different groups of users. Salthouse *et al.* (1978) found that student ratings had virtually no effect upon faculty decisions about promotions and salary increases. Coleman and McKeachie (in preparation) and Borgida and Nisbett (1977) found a small effect of student ratings upon student choice of courses. Both Borgida and Nisbett and the studies reported here indicate that face-to-face communication increases effectiveness.

The research on student ratings over the last half-century has given us a great deal of useful information about the dimensions involved in student ratings, the variables affecting student ratings and the validity of student ratings of instruction. Now, it seems to us, we need to turn to the uses of student ratings for the benefit of, for example, faculty and administrators making decisions about promotions and salaries, faculty members trying to improve their teaching and students trying to choose courses more wisely. How can the information on student ratings be presented so that it will have an impact? If we produce an impact, how can we ensure that it will result in decisions better for education?

The present study demonstrates that consultation can help. Now we need to find out what kinds of consultation are most effective for those being counselled. What kinds of consultation not only affect teacher ratings but also result in greater student learning?

ACKNOWLEDGMENTS.—This research was supported by grant no. NE-G-00-3-0110 from the National Institute of Education.

REFERENCES

- ALEAMONI, L. M. (1974). *The Usefulness of Student Evaluations in Improving College Teaching*. Research Report No. 356. Urbana, Ill.: Measurement and Research Division, Office of Instructional Resources, University of Illinois at Urbana-Champaign.
- BORGIDA, E., and NISBETT, R. E. (1977). The differential impact of abstract vs. concrete information on decisions. *J. appl. soc. Psychol.*, 7, 258-271.
- CARRIER, N. A. (1966). *Evaluating the Introductory Psychology Course*. Reading, Mass.: Addison-Wesley.
- CENTRA, J. A. (1973). Effectiveness of student feedback in modifying college instruction. *J. educ. Psychol.*, 65, 395-401.
- COLEMAN, J., and McKEACHIE, W. J. Effects of instructor/course evaluations on student course selection. (In preparation.)
- GAGE, N. L., RUNKEL, P. L., and CHATTERJEE, B. B. (1963). Changing teacher behavior through feedback from pupils: an application of equilibrium theory. In CHARTERS, W. W., and GAGE, N. L. (Eds.), *Readings in the Social Psychology of Education*. Boston: Allyn and Bacon.
- ISAACSON, R. L., McKEACHIE, W. J., MILHOLLAND, J. E., LIN, Y.-G., HOFELLER, M., BAERWALD, J. W., and ZINN, K. L. (1964). Dimensions of student evaluations of teaching. *J. educ. Psychol.*, 55, 344-551.
- KAHNEMAN, D., and TVERSKY, A. (1973). On the psychology of prediction. *Psychol. Rev.*, 80, 237-251.

- MARSH, H. W., FLEINER, H., and THOMAS, C. S. (1975). Validity and usefulness of student evaluations of instructional quality. *J. educ. Psychol.*, 67, 833-839.
- McKEACHIE, W. J., and LIN, Y-G. (1975). Multiple discriminant analysis of student ratings of college teachers. *J. educ. Res.*, 68, 300-305.
- McKEACHIE, W. J., LIN, Y-G., and MANN, W. (1971). Student ratings of teacher effectiveness: validity studies. *Am. educ. Res. J.*, 8, 435-445.
- MILHOLLAND, J. E. (1964). Measuring cognitive abilities. In McKEACHIE, W. J., ISAACSON, R. L., and MILHOLLAND, J. E., *Research on the Characteristics of Effective College Teaching*. Final Report: Co-operative Research Project No. SAE 850, United States Office of Education and the University of Michigan. Washington, D.C.: US Government Printing Office.
- MILLER, M. T. (1971). Instructor attitudes toward, and their use of student ratings of teachers. *J. educ. Psychol.*, 62, 235-239.
- PAMBOOKIAN, H. S. (1974). Initial level of student evaluation of instruction as a source of influence on instructor change after feedback. *J. educ. Psychol.*, 66, 52-56.
- SALTHOUSE, T. A., McKEACHIE, W. J., and LIN, Y-G. (1978). An experimental investigation of factors affecting university promotion decisions. *J. higher Educ.*, 49, 177-183.
- SULLIVAN, A. M. (1979). The place of research in faculty development. The why—and the how. Paper presented at the Center for Research on Learning and Teaching, the University of Michigan, Ann Arbor, April 5.
- TUCKMAN, B. W., and OLIVER, W. F. (1968). Effectiveness of feedback to teachers as a function of source. *J. educ. Psychol.*, 59, 297-301.

(Manuscript received 26th October, 1979)

UNDERSTANDING THE PURPOSE OF STRATEGY TRAINING

BY M. J. LAWSON AND SANDRA FUELOEP

(*The Flinders University of South Australia*)

SUMMARY. Two groups of retarded adults were trained to use a cumulative rehearsal strategy for a serial recall task. Those subjects who received instructions which emphasised the intended purpose and potential uses of the trained strategy showed significantly better serial recall performance than subjects invited to participate in 'a remembering task'. The superiority in serial recall performance was present up to one week after training ceased. The results are discussed in terms of the role of executive, or metacognitive, skills in the facilitation of transfer.

INTRODUCTION

THE generalisation of learning as assessed by retention and transfer tests is the primary objective of instructional programmes involving the mentally retarded. Attempts to provide evidence of generalisation have largely been unsuccessful and the results of training programmes involving cognitive instruction of the retarded have been reviewed in pessimistic terms (Butterfield *et al.*, 1973; Brown, 1977). In particular, training in the use of mnemonic strategies has not been very successful in producing either long-term retention or transfer of learning to related tasks (Brown, 1977).

Consideration of this situation has led recently to a major reorientation of theoretical approach to the study of both maintenance and transfer of learning in retardates. At a general level this reorientation has been based on the view that it is necessary in training to develop executive, or metacognitive, skills (Butterfield and Belmont, 1975; Brown, 1977). One implication of this view is that memory training studies should be concerned not only with the development of specific acquisition or retrieval strategies but should also address the problems of appropriate use of these strategies by subjects involved in the training. The decision to use a particular strategy is an executive one and would be based on such things as the subjects' understanding of the task, the content to be retained, and the retention interval.

The first of the above-mentioned executive skills, the subjects' understanding of the nature of the task, has generally been neglected in training studies involving mentally retarded subjects. However, the results of several studies with young children indicate that the subjects' assessment of the nature of the task has an important influence on performance. Kennedy and Miller (1976) have shown that the act of pointing out to the child the potential advantages which attend a particular strategy facilitates transfer of that strategy. Borkowski *et al.* (1976) used a film to sensitise first graders to the value of a strategy for learning paired associates and their results suggest that this increased awareness of the value of the strategy increased the use of that strategy in transfer tests.

The deliberate attempts to make the subjects aware of the purpose of training in these two studies may be contrasted with the general instructions in other studies which omit such explanations. In these latter studies it is not uncommon for subjects to be asked to participate in a 'memory game', for example (Neimark *et al.*, 1971; Allik and Siegel, 1976; Brown and Barclay, 1976; Bray *et al.*, 1977). If subjects then treat the experimental task in a game-like fashion, and perhaps assume that it is a game with particular and unique rules, then it is likely that they would treat a transfer task as a new 'game' with another set of rules. If this is the case then both maintenance and transfer of a trained strategy is likely to be inhibited.

In this study, which forms part of a larger study of the executive skills of retarded adults, we were concerned with assessing the effects on retention of instructions which provided subjects with differing degrees of information about the purpose of the strategy training.

METHOD

Subjects

24 retarded adults, 13 females and 11 males, from three workshops in suburban Adelaide, participated in the study. All subjects attended the workshops on a daily basis and were involved in largely light industrial activities. Mean age of subjects was 29.9 years, mean performance IQ was 55.5 and mean period of attendance at the workshop was 9.3 years.

Subjects were assigned to one of two groups. In the Standard Instructions group subjects were invited to—"take part in a remembering task. In this task we will look at some pictures and will put the pictures on this board. I want you to remember the position of each picture on the board." Although these subjects were aware that some form of training was to be undertaken no discussion of the purpose of the training was entered into.

In the Purposive Instruction group, in addition to the above general instructions, subjects were involved in a discussion of the purpose of training prior to the start of the first training session and during each subsequent session. They were told that:

"In this remembering task we will tell you a way which will help you to remember better. We will give a rule and this rule will help you to remember more and more pictures. You can use the rule to help you remember other lists of things."

Both groups were given the same strategy training. The discussion of the purpose of training constituted the difference in treatment.

Stimuli

Stimuli were coloured illustrations of common objects taken from a children's picture dictionary and mounted on 6 × 6 cm cards. 35 mm slides of each card were also prepared. The cards and slides were arranged into two groups of lists with list length in each group ranging from two to 14 items. No two items in any list were obvious members of the same category and care was taken to separate items which might be easily associated on the basis of sound, shape, or function. Both groups of lists were presented to each subject with order of presentation of items within a list being changed for successive training sessions.

Strategy training

During training subjects were trained in the use of a cumulative serial-recall strategy based on that used by Palmer and Ornstein (1971). This involved:

- (a) Three repetitions of the name of the first item in the list, e.g., "Shoe, shoe, shoe".
- (b) Following placement of the second item, three repetitions of both the items, e.g., "Shoe-knife, shoe-knife, shoe-knife".
- (c) Cumulative repetition of all items in the same manner until the end of the list was reached.

Assistance was given by the experimenter in the form of a reminder to continue use of the strategy for each item.

Subjects were tested individually in small offices in the workshops. In an initial session subjects were asked to label the stimulus items. Labels were provided if items could not be named. Subjects were also given a serial recall task involving

objects collected from around the workshop in order to ensure that the nature of the recall task was clearly understood.

In training sessions a slide of each item was first projected on a screen for three seconds. Subjects were then presented with a card illustration of that item and were required to place that card face-up in the left-hand slot on the top row of a wooden display board. Once subjects had completed the strategy instruction for that card it was turned face down. The display board had 14 slots on each of its top and bottom rows.

After the presentation of the last slide in the list and the placement of the corresponding card the subjects were given a duplicate set of list items and were asked to place these in the bottom row of the display board in the same order as that on the row above. For this recall task the duplicate cards were set out in random order in a roughly circular fashion on a desk in front of the subject. The period between placement of the last card in the top row and placement of the first duplicate card was approximately 20 seconds. Subjects were encouraged to rehearse during this interval.

Once the last card had been placed in the bottom row of the display board the cards in the top row were turned around and the two orders of cards compared. Feedback information from this comparison was discussed with subjects and was made more concrete by use of a felt display board on which green and white circles of felt could be arranged. A row of green circles was placed on this board corresponding to the items on the top row of the wooden display board. Comparisons were then made between top and bottom rows, green circles being placed on the bottom row where the orders of the two rows corresponded, white circles being used to indicate inaccurate placements.

Each experimental session lasted approximately 45 minutes. Subjects received either seven training sessions or continued the training procedure until the same level of performance was maintained over two consecutive days. One subject attended six sessions and five subjects needed eight sessions to satisfy the above criteria. (The shortest list recalled successfully was seven items in length, the longest 13 items. The profile for the best performance was similar for both groups.)

Lists were presented in ascending order and training in each session continued until a subject failed to recall a list of a given length on two consecutive trials. The length of the longest list placed correctly in presentation order was taken as the best level of performance (Best).

Prior to training in use of the strategy all subjects were allowed to complete the task on their own. Performance on this task was taken as the pre-training score (Pretraining).

In the test phase of the study three retention tests were administered in the following order:

- (1) *Hour*: In the session following attainment of the best level of performance subjects were given a short strategy training session. One hour later they were given the serial-recall task without any assistance (in terms of a task reminder) from the experimenter.
- (2) *Week*: One week after the above session the same task was given to subjects again without assistance from the experimenter.
- (3) *Week-reminder*: Following the attainment of the best level of performance without assistance, subjects were shown their levels of performance with assistance (Best) and without assistance (Week) and were asked by the experimenter to do the task 'the way that you did it when I was helping you'. No details of the strategy were given at this time. This condition

was included to provide some further information about strategy use, i.e., to see whether or not any decrement in serial recall performance over time could be attributed to failure to remember and use the trained strategy.

RESULTS

The length of the longest list correctly placed in order of presentation was the dependent variable in the analysis. The mean number of items placed correctly for each test occasion is shown in Table 1. The design was a 2×5 (Group $\times$ Test Occasion) factorial with repeated measures on the second factor.

The main effect for Group failed to reach significance $F(1, 22) = 3.99$, $P < 0.06$. Both the main effect for Test Occasion, $F(4, 88) = 72.9$, $P < 0.001$, $MSe = 1.13$, and the Group $\times$ Test Occasion interaction, $F(4, 88) = 5.97$, $P < 0.001$, $MSe = 1.13$, were significant. The main effect for Test Occasion reflects the fact that subjects Best scores were significantly higher than those for the other occasions, and that their scores on the three retention tests were significantly better than the Pre-training scores.

TABLE 1
MEAN LENGTH OF LONGEST LIST CORRECTLY PLACED IN PRESENTATION ORDER

Group	Test occasion					Mean
	Pre- Training	Best	Hour	Week	Week- reminder	
Standard Instructions (N = 12)	4.42	9.08	5.67	5.58	5.83	6.12
Purposive Instructions (N = 12)	4.17	9.83	7.92	6.92	8.08	7.38
Mean	4.29	9.46	6.79	6.25	6.96	

The result of major interest for this study is the significant Group $\times$ Test Occasion interaction. Subsequent analysis of this interaction by means of a simple effects analysis (Winer, 1962, p. 310), showed that the effect of the Purposive Instructions was significant for both the Hour retention test, $F(1, 22) = 9.92$, $P < 0.01$, $MSe = 3.32$, and for the Week-reminder conditions, $F(1, 22) = 9.14$, $P < 0.01$, $MSe = 3.32$. There was no comparable effect at any of the other three test occasions. For the Week retention test, although the scores for the Purposive Instructions group were higher than those of the other group, they were not significantly so, $F(1, 22) = 3.21$, $P > 0.05$. The same pattern of results was obtained when pairs of means were compared at each level of the Test Occasion factor, using the Tukey multiple comparison procedure (Keppel, 1973, p. 245).

DISCUSSION

These results are of interest at several levels. The pattern of performance of the total group conforms to that present in other mnemonic-strategy training studies (see Brown, 1977), a pattern which is held to be indicative of a production deficiency. Retarded subjects can learn to use an acquisition strategy under instruction but fail to maintain that strategy when assistance provided by constant instruction is withdrawn. In this study the above pattern is shown by the decrement in performance following the withdrawal of assistance after the Best scores were attained.

The beneficial effect of the reminder in the Week-reminder reinforces this production-deficiency interpretation. While the effect of the Purposive Instructions

is strong, it is apparent that the degree of retention of the trained strategy, particularly by subjects in the Standard Instructions group, was relatively small. Data which would provide an explanation of the reminder effect are not available in this study though it seems likely that the reminder did lead to reinstatement of the trained strategy for some subjects. This view is supported by the fact that 11 of the 24 subjects did increase their serial recall scores when given the reminder. However, as is apparent when the means of the two groups are compared for the Week-reminder occasion, the majority of these subjects (9 of 11) were in the Purposive Instructions group.

The focus of this study was the effect of the Purposive Instructions on maintenance of a trained strategy, and this effect is represented in the Group $\times$ Test Occasion interaction. Although the groups did not differ on Pre-training or Best scores the Purposive Instruction group did show a consistent superiority over subjects in the Standard Instructions group on each of the three retention tests. Instructions which stressed the purposive nature of the training, the rule which was to be used, and the potential uses of this rule, did result in significantly better serial-recall performance for at least one week after the cessation of training. Although from this study it is not possible to determine which specific component, or components, of these instructions is responsible for the facilitative effect, it is clear that discussion with subjects of the nature and intended uses of a strategy for serial recall does result in significantly better utilisation of that strategy. These results obtained with a retardate population parallel those established by Borkowski *et al.* (1976) and Kennedy and Miller (1976) in their studies with young children.

At a more general level the results suggest that the likelihood of transfer, as well as maintenance, of a trained strategy will be increased by increasing the subjects' awareness of the purpose of training. Without some understanding of the experimenter's purpose in providing training it is likely that subjects become involved in what Brown (1977) describes as 'a blind-rule-following' exercise. The effect of such misattribution of purpose would seem to be most powerful in training situations which are represented as 'remembering games'. In these contexts transfer of learned rules by mentally retarded subjects may well be inhibited.

ACKNOWLEDGMENTS.—This research was supported by a grant from the Australian Research Grants Committee. The authors wish to acknowledge the assistance and co-operation of Mrs. Elaine White and the staff of the three workshops of the Mentally Retarded Children's Society of South Australia. Requests for reprints should be addressed to Michael J. Lawson, School of Education, Flinders University, Bedford Park, 5042, South Australia.

REFERENCES

- ALLIK, J. P., and SIEGEL, A. W. (1976). The use of the cumulative rehearsal strategy: a developmental study. *J. exper. Child Psychol.*, 21, 316-327.
- BORKOWSKI, J. G., LEVERS, S., and GRUENEFELDER, T. G. (1976). Transfer of mediational strategies in children: the role of activity and awareness during strategy acquisition. *Child Dev.*, 47, 779-786.
- BRAY, N. W., JUSTICE, E. M., FERGUSON, R. P., and SIMON, D. (1977). Developmental changes in the effects of instructions on production-deficient children. *Child Dev.*, 48, 1019-1026.
- BROWN, A. L. (1977). Knowing when, where and how to remember: a problem of metacognition. *Technical Report No. 47*. Centre for the Study of Reading. University of Illinois at Urbana, Champaign.
- BROWN, A. L., and BARCLAY, C. R. (1976). The effects of training specific mnemonics on the metamnemonic efficiency of retarded children. *Child Dev.*, 47, 71-80.
- BUTTERFIELD, E. C., and BELMONT, J. M. (1975). Assessing and improving the cognitive functions of mentally retarded people. In BAILER, I., and STERNLICHT, M. (Eds.), *Psychological Issues in Mental Retardation*. Chicago: Aldine.
- BUTTERFIELD, E. C., WAMBOLD, C., and BELMONT, J. M. (1973). On the theory and practice of improving short-term memory. *Am. J. ment. Defic.*, 77, 654-669.
- KENNEDY, B. A., and MILLER, D. (1976). Persistent use of verbal rehearsal as a function of information about its value. *Child Dev.*, 47, 566-569.

RESEARCH NOTES

Br. J. educ. Psychol., 50, 181-185, 1980

THE ATTITUDES OF TEACHERS IN FURTHER EDUCATION

By G. M. YOUNG

(School of Education, University of Leicester)

SUMMARY. On the basis of earlier research in the primary and secondary sectors, three dimensions of educational attitude were hypothesised to exist among teachers in further education. Principal components analysis of the scores on a questionnaire of teachers and tutors in this sector uncovered three factors resembling the characteristics of the predicted constructs.

INTRODUCTION

Traditionally, the search for common sources of variance in the educational attitudes of teachers has been expressed in multi-dimensional terms in the pursuit of a number of separate and orthogonal factorial dimensions. Oliver (1953), indeed, suggested that even the most economical account of the polarities of attitude on educational issues might require at least three orthogonal dimensions carrying six oppositional systems of attitudinal belief. Factor analysis by Butcher (1959) helped sharpen and refine the scales developed by Oliver to measure his predicted dimensions. The resulting Manchester Survey of Opinions about Education (Oliver and Butcher, 1962) has been widely used to study the attitudes of primary and secondary teachers for whom its item content was clearly constructed. What has been lacking until now has been a similar exploration in depth of the attitudes of teachers in the further education sector. It was with such a project in mind that the present research was mounted. Three domains of attitudinal issues of assumed relevance to teachers in colleges of further education were chosen as targets for analysis: pedagogy and teacher-student interaction; the distribution of access to the resources of educational systems; the curriculum. Each was hypothesised to be the focus of an evaluative dimension carrying two opposed systems of educational attitude.

The hypothesised dimensions

(i) Naturalism-Formalism in attitude to pedagogic role and teacher-student interaction

The Naturalism-Formalism dimension carries an opposition of attitudes that previous research has found to pervade the secondary sector. Thus 'Naturalism' was one of the constructs hypothesised by Oliver (1953), which McLeish (1969) found to be negatively related ($r: -0.40$) to Steele's (1958) Formalism scale. Naturalism favours an active, exploratory, and autonomous role for the student in pursuit of the harmonious development of his idealised inner 'needs' whereas Formalism is perhaps best encapsulated by Steele (1958) as a tendency to value standards of achievement and good behaviour in a teacher-centred classroom where planned activities are conducted in a systematically ordered programme of studies.

(ii) Egalitarianism-Conservatism in attitude to the distribution of access to educational resources

Egalitarianism-Conservatism represents a transposition to the further education sector of many of the attitudes of Oliver's (1953) Radicalism scale but with a sharper focus upon the value of social equality. Egalitarianism favours the more equitable redistribution of collective access to the material and symbolic resources of a stratified society. In opposition to Egalitarianism at the other end of this resources dimension is a cluster of attitudes that favour individual competition for the limited resources of an unchanging system. Whereas at the egalitarian pole the pools of distributable wealth, 'talent', or 'educability' are seen as capable of almost limitless growth, the conservative pole tends to attract static and closed models of resources and pessimistic attitudes about their more equitable redistribution through programmes of institutional reform.

(iii) Liberalism-Utilitarianism in attitude to the curriculum

The dilemma faced by the attitudinal polarities of the Liberal-Utilitarian dimension is that of whether the goals of the curriculum are to be regarded as values which are intrinsic or extrinsic to the process of schooling. Connecting the attitudes of the Liberal polarity is

the central belief that the values intrinsic to the educational process are of a universality to warrant its almost total liberation from external constraints. The attitudes connected are those of opposition to the demands that threaten constraint—demands of economic and social 'need', of the useful, the vocational, the specialised, the publicly assessable and the applied, and their embodiment in client pressure for control over the curriculum. Opposing Liberalism is an attitudinal orientation that would evaluate the outcomes of schooling by the criterion of their instrumentality to the achievement of national power, order, and industrial efficiency. Precursors of this Liberal-Utilitarian dichotomy are to be found in constructs hypothesised in earlier researches such as Gamson's (1966) dimension of Normative vs. Utilitarian orientations to education, Healey and Holloway's (1968) conception of Liberal and Utilitarian objectives in higher education, and Oliver's (1953) dimension of Tough-Tendermindedness in Education redefined by Butcher (1959) as one of conflict between cultural and vocational goals in the process of schooling.

(iv) *Operational definition of the hypotheses*

It was hypothesised that a principal component analysis of the scores on a questionnaire on the educational beliefs of teachers and tutors in further education would reveal three separate factors of response interpretable by their significant loadings as the attitudinal dimensions of Naturalism-Formalism, Egalitarianism-Conservatism and Liberalism-Utilitarianism in Education.

METHOD

The sample

Since heterogeneity is a more critical feature than representativeness in this kind of research, the aim of the sampling design was less concerned with the latter than with ensuring that respondents reflected the diversity of personnel in the sector of further education. As researchers in other sectors (Pollock, 1964; Shipman, 1965; McLeish, 1970) had found contrasts in attitude between tutors in education and practising teachers, to ensure such diversity at the attitudinal level it was decided to sample not only the teaching personnel of the colleges of further education but the tutors of the colleges of education (technical) mainly responsible at the time the research began for the training of teachers in further education. The colleges of education (technical) were sampled on a national basis and the more multitudinous colleges of further education on a regional basis. The response from the 10 co-operating colleges (three national colleges of education (technical) and seven Lancashire colleges of further education) produced a sample large by the norms of this kind of study: 427 tutors and teachers concerned with the sector of further education.

The questionnaire

The instrument of measurement was a Likert-type questionnaire containing 210 statements of belief chosen to reflect the dual attitudinal orientation of each predicted dimension. Response to each item was scored for endorsement or rejection along a conventional five-point scale of intensity with a neutral midpoint. The document carried three separate sections containing each a pool of items expressive of one of the predicted bipolar constructs and a fourth of items concerned with commitment to teaching as a profession. To assist in the clarification of the dimensions, 13 'marker' or criterion variables were included in the instrument: six items (two from Steele's (1958) Formalism measure, two from Kissack's (1956) measure of attitude to corporal punishment and two from Oliver and Butcher's (1962) Naturalism scale) to help identify the Naturalism-Formalism dimension; four (two from Eysenck's (1944, 1947) and two from Oliver and Butcher's (1962) Radication scales) to identify the dimension of Egalitarianism-Conservatism; and finally three (two from Higson's (1951) measure of educational interests and one from Oliver and Butcher's (1962) Tough-Tendermindedness scale) to help in the identification of the Liberal-Utilitarian dimension made up the 13. To reduce the matrix to a manageable size for the computer, the final document of 210 items was thinned out by removing non-discriminating items by traditional methods of item analysis (Edwards, 1954). 97 items and 8 criterion variables survived the item analysis and were subjected to principal components analysis with rotation to the varimax (Kaiser, 1958) and promax solutions. The analysis of the factor matrices was eventually terminated at the point of closest approximation between the saturations of the extracted factors and the characteristics attributed to the hypothesised constructs.

FIGURE 1

THE HIGHEST LOADING ITEMS OF THE EMERGING DIMENSIONS

FACTOR 2ND THIRD ORDER VARIMAX FACTOR II: NATURALISM-FORMALISM			FACTOR 2ND UNROTATED THIRD ORDER COMPONENT II: EGALITARIANISM-CONSERVATISM			FACTOR 4TH THIRD ORDER VARIMAX FACTOR IV LIBERALISM-UTILITARIANISM		
FOCUS: PEDAGOGY & TEACHER- STUDENT INTERACTION			FOCUS: SYSTEMS AND RESOURCES			FOCUS: THE CURRICULUM		
ITEM STATEMENT	ITEM NO.	ITEM LOADING	ITEM STATEMENT	ITEM NO.	ITEM LOADING	ITEM STATEMENT	ITEM NO.	ITEM LOADING
Too much of student effort in F.E. is expended in copying notes from the blackboard.	N8	-.539	The first concern of any wise educational policy should be to find and train a natural elite.	E 60	+.393	There should be less emphasis in education on the acquisition of knowledge and skills and more on developing a willingness to learn.	N7	-.386
Students should not be allowed to wander around the college premises at will during their free periods.	N17	-.538	Education already takes up too high a proportion of public expenditure.	E 53	+.351	If a subject can be taught it can be programmed.	L102	-.375
In the classroom teachers should listen more and talk less.	N11	-.534	Courses in F.E. should be made more directly relevant to the immediate needs of particular employees.	P 77	+.340	A course of instruction in a works school can be as truly educational as a course in a technical college.	L75	-.359
Too often teachers dismiss as an act of insubordination what is really a 'cry for help'.	N12	-.526	The classroom is not a suitable place for the discussion of controversial political issues.	L 97	+.315	Only through a training in the humanities can a technologist be given an understanding of the art of managing people.	L93	-.344
Students in schools & colleges are subject to too many petty rules & regulations.	N5	-.524	The sociological study of education should be given a more prominent place in courses of teacher training.	E 39	+.310	There is too much emphasis in F.E. on the acquisition of technical knowledge & skills.	L83	-.328
All classroom doors should be locked between lessons.	N22	-.503	The taxpayers have more right to interfere in the running of universities than the students whose education they finance.	E 58	+.306	The proper study of mankind is man.	L91	-.309
						The bias towards the applied sciences in F.E. should be strictly maintained.	L96	-.285
Student interest can best be maintained in an informal classroom situation in which there is opportunity to challenge & discuss	N10	-.455	An unemployed school leaver should automatically have the right to enter some form of vocational training in a technical college.	E35	+.305	Too much of student effort in F.E. is expended in copying notes from the blackboard.	N8	-.275
There should be less emphasis in education on the transmission of knowledge & skills & more on developing a willingness to learn.	N7	-.412	Less able adolescents should not be encouraged to develop high academic ambitions.	E37	+.303	Day-release to a local technical college should be made compulsory for all young people under 18 not in full-time education elsewhere.	E37	-.274
Schools and colleges should be more willing to adapt their work to meet the special needs of immigrant students	E36	-.401	Students should be given substantial representation on the governing bodies of colleges.	E44	+.298	Junior teaching staff should be given a greater share in determining the academic policy of their establishment.	P67	-.265
Teachers should not be expected to modify their teaching methods to meet the needs of individual students.	N18	-.366	The primary concern of teachers in F.E. should be with personal development of the individual rather than his job effectiveness.	L80	+.294	In education the literary & the aesthetic rarely receive the attention they deserve.	L87	-.263
It is not the teachers responsibility to motivate the students.	N19	-.343						
PERCENTAGE OF VARIANCE:		5.144	PERCENTAGE OF VARIANCE:		3.132	PERCENTAGE OF VARIANCE:		2.461
						N: Items predicted to load on the Naturalism-Formalism dimension.		
						E: Items predicted to load on the Egalitarianism-Conservatism dimension.		
						L: Items predicted to load on the Liberalism-Utilitarianism dimension.		

RESULTS

The emergence of the Naturalism-Formalism dimension

The failure of a large Naturalism factor to appear among the first-order components compelled the pursuit of the first hypothesis in the second-order structure. Oblique rotation of the second-order axes detected the central cluster of Naturalism-Formalism items that had been dispersed among the first-order components. It was this central cluster of items that saturated the two factors that confirmed the hypothesis in third-order structure in the shape of the strongest unrotated component (accounting for 10.875 per cent of the variance) reappearing in the pattern of Figure 1 as a varimax factor (accounting for 5.14 per cent of the variance) with 12 of its 14 highest loadings and two of its validating criterion items from the pool of statements from the Naturalism section.

The emergence of the Egalitarianism-Conservatism dimension

The pool of Egalitarianism-Conservatism items was also dispersed among a number of weak and narrowly focused factors in the first-order structure. A wider factor of these items emerged among the second-order components, surviving both orthogonal and oblique rotation and re-emerging to confirm the hypothesis in the shape of the second third-order component (accounting for 3.132 per cent of the variance) of Figure 1 which reappeared like the preceding dimension as a varimax factor (accounting for 4.497 per cent of the variance) carrying one of the criterion variables from the Egalitarian pool.

The emergence of the Liberalism-Utilitarianism dimension

Items from the third of the sections also suffered a dispersal among weak and narrowly focused factors in first-order structure. It was items from two of the latter that formed the heaviest saturations of the second-order factor that resurfaced in wider form (but with some extraneous items) to confirm the third hypothesis as the weakest of the third-order factors with only 2.461 per cent of the variance but with 9 of its 16 highest loadings from the section of Liberalism-Utilitarianism items.

DISCUSSION

In Figure 1 the statements chosen to express three dualities of attitudinal orientation can be seen emerging along their predicted factorial dimensions. We have seen how each of the latter connects with a line of content delineation from earlier studies in a way that gives it a tentative claim to factorial invariance across a generation of research. The extraction of a Naturalism-Formalism factor unites the constructs of Naturalism (Oliver, 1953) and Formalism (Steele, 1958) developed with different sectors of teachers in earlier research. A similar claim to invariance has been made both for a dimension of institutional reform in the shape of the Egalitarianism factor and for one of conflict between the cultural and vocational objectives of schooling. Moreover, each of these dimensional lines of research has now been extended to a little researched area of the educational world, that of teachers in the sector of further education. Nor were the factors extracted mere surface features to be reduced to a deeper, more parsimonious construct. They were themselves a condensation of second-order structure and as third-order factors not reducible further to any meaningful degree. The patterns uncovered might thus provide some useful co-ordinates for future research on the attitudes of teachers in this little known sector of the educational system.

REFERENCES

- BUTCHER, H. J. (1959). *The Opinions of Teachers and Student Teachers about Education*. University of Manchester: unpublished Ph.D. Thesis.
- EDWARDS, A. L. (1954). *Techniques of Attitude Scale Construction*. New York: Appleton-Century-Crofts.
- EYSENCK, H. J. (1944). General social attitudes. *J. soc. Psychol.*, 19, 207-227.
- EYSENCK, H. J. (1947). *Dimensions of Personality*. London: Routledge and Kegan Paul.
- GAMSON, Z. F. (1966). Utilitarian and normative orientations towards education. *Sociol. Educ.*, Vol. XXXIV, 54, 46-47.
- HEALEY, F. G., and HOLLOWAY, P. J. (1968). The 'classical' objectives of university teaching. In BEARD, R. M., HEALEY, F. G., and HOLLOWAY, P. J., *Objectives in Higher Education*. S.R.H.E. 24-40.
- HIGSON, F. G. (1951). *An Enquiry into the Interests in Education of Teachers in Training*. University of Manchester: Unpublished M.Ed. Thesis.

- KAISER, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- KISSACK, M. (1956). The Attitudes of Training College Students towards Corporal Punishment. University of Manchester: Unpublished M.Ed. Thesis.
- MCLEISH, J. (1969). *Teachers' Attitudes: a Study of National and Other Differences*. Cambridge Institute of Education.
- MCLEISH, J. (1970). *Students' Attitudes and College Environments*. Cambridge. Institute of Education.
- OLIVER, R. A. C. (1953). Attitudes to education. *Br. J. educ. Studies*, 2, 31-41.
- OLIVER, R. A. C., and BUTCHER, H. J. (1962). Teachers' attitudes to education, etc. *Br. J. soc. clin. Psychol.*, 1, 32-37.
- POLLOCK, G. J. (1964). The Opinions of Science Teachers on the Objectives of Teaching Science. University of Manchester: Unpublished M.Ed. Thesis.
- SHIPMAN, M. (1965). Personal and Social Influences on the Work of a Training College. University of London: Unpublished Ph.D. Thesis.
- STEELE, P. W. (1958). Change in Attitude amongst Training College Students towards Education in Junior Schools. University of Manchester: Unpublished M.Ed. Thesis.

(Manuscript received 22nd August, 1979)

Br. J. educ. Psychol., 50, 186-187, 1980

A TRANSMISSION-INTERPRETATION SCALE

By P. L. GARDNER AND S. M. TAYLOR

(Faculty of Education, Monash University, Australia)

SUMMARY. Barnes and Shemilt (1974) have conceptualised transmission-interpretation (T-I) as an important dimension of teachers' verbal behaviour. Transmissive teachers are held to differ from interpretative ones in their beliefs about the nature of knowledge, in their expectations of student performance and in their perceptions of their tasks as teachers. These differences in beliefs, expectations and perceptions might be expected to surface as observable differences in the way teachers use language in the classroom.

The present paper describes an attempt to measure teacher T-I by means of a 40-item Likert scale completed by students. After correction for attenuation, the T-I scale correlated 0.83 with the Student Perception of Teacher Style scale (Tuckman, 1970) which measures teacher directiveness. A shortened (20-item) version of the T-I scale yielded an estimated coefficient alpha of 0.93, suggesting that the construct can be measured with a high degree of internal consistency.

INTRODUCTION

Barnes and Shemilt (1974) have conceptualised transmission-interpretation (T-I) as an important dimension of teachers' verbal behaviour. The *transmissive* teacher is thought of as one who: (a) regards knowledge as the content of the various academic disciplines; (b) values students' performances to the extent that they match the criteria of the disciplines; (c) sees the task as the evaluation and correction of the learner's performance; (d) considers students as generally uninformed people without much relevant knowledge.

The *interpretative* teacher, on the other hand: (a) regards knowledge as a personal ability to organise thoughts and actions; (b) values students' own interpretations of their observations; (c) sees the task as the promotion of discussion through which personal knowledge can be reshaped; (d) considers that learners possess relevant knowledge which can be drawn upon.

METHOD OF MEASUREMENT

In order to measure teacher transmissiveness, a 20-item five-point Likert scale was developed. The scale focuses upon aspects of the verbal behaviour of teachers in the classroom, and draws upon student perceptions of that behaviour. Student ratings were employed for several reasons. First, if teacher behaviour *as perceived by students* (rather than teacher behaviour *per se*) influences student outcomes, then student ratings are potentially more valid for research on teacher effects. Second, they are potentially more reliable, because of the extended period of participant observation, and because the ratings of many students can be averaged to provide an estimate of the consensus of the class. Finally, they are easy and economical to gather and analyse.

RESULTS AND DISCUSSION

Scale development

An initial pool of 37 items was constructed. The content validity of these items was judged by a panel of 20 faculty members and graduate students with interests in psychology, measurement and/or teaching behaviour. Revisions and additions resulted in a trial form containing 40 items.

The trial form was administered to 362 Year 10 students in 15 different classes in 14 co-educational high schools located in middle-class suburbs of the Melbourne metropolitan area. Data were gathered on 53 teachers. The trial data were subjected to item analysis and factor analysis (both minimum residual and varimax). The factor analyses were carried out on the matrix of 53 teachers $\times$ 40 class mean item scores. In a conventional application of factor analysis, the matrix of 362 students $\times$ 40 individual item scores might have been analysed, but this would have resulted in data which would have been difficult to interpret, primarily because item variance would partly have been due to variations among teachers, and partly to variations in perception among students within classrooms. Analysis of class mean item scores reduces the latter source of variance and clarifies the resultant factor pattern.

Final form

A procedure devised by McKennell (1970) was then used to produce a final form of the T-I scale. The procedure, which utilises item intercorrelations and Cronbach's alpha coefficient, maximises the internal consistency of a scale, by setting off the advantage of increasing the number of items in the scale against the disadvantage of having to include more and more items of decreasing average inter-item correlation. The estimated Cronbach alpha coefficient became a maximum of 0.93 for 20 items. These items are presented in Appendix 1. All these items have item/rest-of-test correlations in excess of 0.30, and all load substantially (mostly 0.4 to 0.7) on one factor.

Construct validity

Students in the trial sample were also asked to complete the *Student Perception of Teacher Style* (SPOTS) scale (Tuckman, 1970), a 17-item 9-point rating scale designed to assess teacher direct/indirect influence (Flanders, 1970). The correlation between the 40-item trial T-I scale and SPOTS was 0.62, which becomes 0.83 when corrected for the unreliability of both measures (Thorndike and Hagen, 1969, p. 198). T-I therefore appears to have much in common with direct/indirect teacher behaviour; the T-I scale, however, appears to be a more reliable measure of the construct than the SPOTS. Further work to validate the T-I, and to use it in substantive research, is currently under way.

REFERENCES

- BARNES, D., and SHEMILT, D. (1974). Transmission and interpretation. *Educ. Rev.*, 26, 213-228.
 FLANDERS, N. A. (1970). *Analyzing Teaching Behavior*. Reading, Mass.: Addison-Wesley.
 MCKENNEL, A. (1970). Attitude measurement: use of coefficient alpha with cluster or factor analysis. *Sociology*, 4, 227-245.
 THORNDIKE, R. L., and HAGEN, E. P. (1969). *Measurement and Evaluation in Psychology and Education*. (3rd edit.) New York: Wiley.
 TUCKMAN, B. W. (1970). A technique for the assessment of teacher directiveness. *J. educ. Res.*, 63, 395-400.

(Manuscript received 22nd August, 1979)

APPENDIX 1

TRANSMISSION-INTERPRETATION SCALE

- 1 (T) In most lessons, my teacher does nearly all the talking.
- 2 (I) Our teacher often helps us to connect new ideas with ideas we have learned previously.
- 3 (I) In this class, the teacher allows us to talk about new ideas with other students.
- 4 (T) Our teacher spends a lot of time just following the textbook.
- 5 (T) Our teacher cuts students off in mid-sentence if their answers aren't along the right lines.
- 6 (T) Much of what we learn in this class is irrelevant to the world outside the classroom.
- 7 (I) When we meet a new idea, our teacher explains the evidence and reasoning very carefully.
- 8 (T) Our teacher usually asks the bright students in the class to answer questions because they are more likely to know the right answers.
- 9 (T) Discussion of alternative ways of solving a problem doesn't happen very often in this class.
- 10 (I) Our teacher often says something like "You'd better learn this thoroughly because it's going to be in the exam".
- 11 (I) The teacher allows us to break up into small discussion groups to talk over an interesting idea.
- 12 (T) If a student gives a wrong answer, the teacher often ignores it or says "no" and then asks someone else.
- 13 (T) My teacher likes to get on with the syllabus.
- 14 (I) Our teacher values students' opinions and ideas about a topic.
- 15 (T) My teacher is more interested in checking that we know the right answers, than in our thinking and reasoning.
- 16 (T) In most lessons, our teacher talks and we just sit and listen, or take notes.
- 17 (I) In this class we are given a chance to talk about our experiences if they are related to the lesson topic.
- 18 (I) Our teacher really listens when a student is giving an answer or an opinion.
- 19 (T) Our teacher can get annoyed if someone doesn't know the right answer to a question.
- 20 (I) In this class, students are encouraged to ask the teacher questions.

NOTE

- (T) Transmissive item: scored Strongly Agree = 5, Agree = 4, Disagree = 2, Strongly Disagree = 1, Not sure = 3.
 (I) Interpretation item: scored Strongly Agree = 1, Agree = 2, Disagree = 4, Strongly Disagree = 5, Not sure = 3.

CONFLICT BETWEEN CUES IN NUMBER CONSERVATION TASKS

BY CAROLE MILLAR AND C. K. MACKAY

(Department of Psychology, University of Aberdeen)

SUMMARY. Thirty primary school children, aged between 5 years 0 months and 5 years 11 months, were given number conservation tasks adapted from Bryant (1974). The tasks were at two levels, viz., within the children's subitising range and within their counting range. For the experimental group, elements of the two rows were joined by thread thus provoking a one-to-one correspondence across rows. The experimental group's performance was very significantly better than that of the control group thus supporting the view that, when one-to-one correspondence is a salient cue, young children can apply the principle of invariance effectively. A significant difference between performance at the subitising and counting levels is interpreted as support for the view that number invariance is initially developed for small numbers and then extended into the counting range.

INTRODUCTION

In the traditional number conservation experiments, conservation is said to occur when the number of elements is judged constant despite transformations in the disposition of these elements. Piaget (1952, 1968) believes that conservation is dependent on the attainment of mental operations and, at the level of operatory conservation, children judge quality or inequality of two arrays by looking for a one-to-one correspondence between their elements irrespective of the disposition of the elements. At the pre-operational stage, children are likely to judge one collection to be more numerous than the other when the collections are presented as rows of different length. At the transitional stage, children strive to use the one-to-one correspondence cue but will adopt a length cue when that cue is exaggerated.

Bryant (1974) has proposed a 'conflict hypothesis' to account for failure at the transitional stage. He suggests that young children experience conflict between a 'perceptual' cue (length) and a conceptual one (the principle of invariance) and that failure to maintain quantitative relations across changes of configuration is due to the inability of the young child to recognise that the one-to-one correspondence cue is a reliably correct one while the length cue is not. Training that length is a 'bad' cue and one-to-one correspondence is a 'good' cue by giving children feedback on their judgments of more and less numerous under different conditions was effective in increasing their dependence on one-to-one correspondence judgments (Bryant 1974, chapter 8).

Both Piaget and Bryant seem to assume that the one-to-one correspondence cue is the preferred one although neither appears to have tested that assumption. In the present note, an attempt is made to do so.

An attempt is also made to examine the claim made by Young and McPherson (1976) that children exhibit an understanding of the principle of invariance with respect to small numbers before they show the same understanding with respect to larger numbers. Klahr and Wallace (1973) distinguished three basic systems used to generate quantitative symbols. These 'Quantification operators' they termed the subitising operator (responsible for the rapid processing of small numbers), the counting operator (responsible for the enumeration of numbers up to about 30), and the estimation operator (responsible for the evaluation of large numbers and continuous quantities). The hypothesis generated from the Young and McPherson study is that children will demonstrate an understanding of the principle of invariance in dealing with numbers within their subitising range before they do so for larger numbers within their counting range.

METHOD

Subjects

The subjects were 30 children from a primary school in Aberdeen. Their ages ranged from 5 years 0 months to 5 years 11 months with a mean C.A. of 5 years 5 months. Children were allocated to the control or experimental group to ensure equivalence of age and sex distribution. There were 9 boys and 6 girls in the control group and 8 boys and 7 girls in the experimental group.

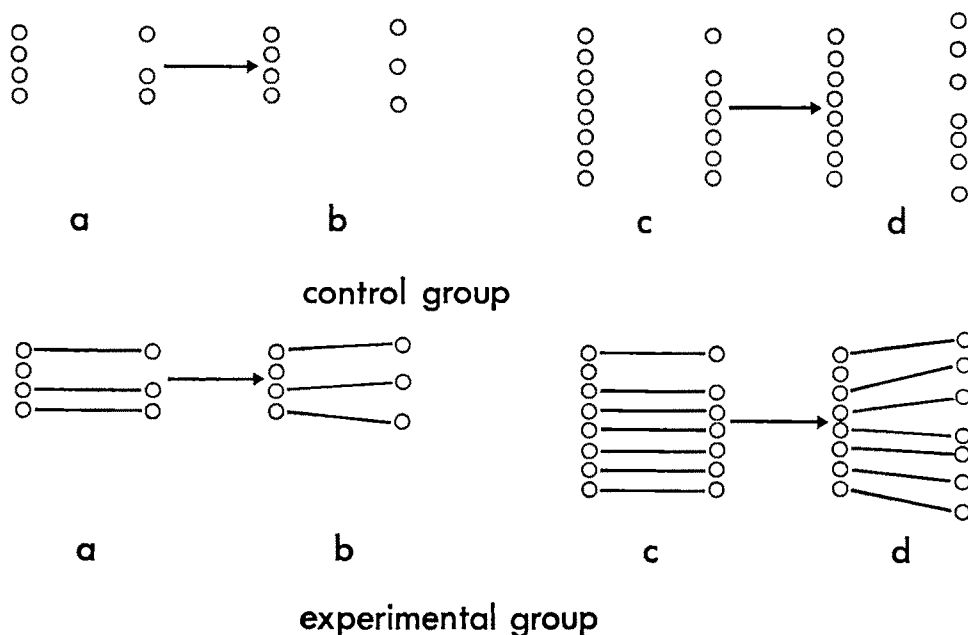
Procedure

Each child was seen individually. Before being asked to attempt the conservation task, it was ascertained whether the child could count to at least eight and whether he/she could subitise beyond two. Counting was assessed by asking the child to count aloud a row of 12 counters. Subitising level was assessed using the procedure described by Young and McPherson (1976). Seven plain white cards, each measuring 15×10 cm, had black dots of 0.8 cm diameter printed on them. The number of dots, arranged randomly, on each card ranged from one to seven. A card was covered by the experimenter's hand and presented to the child. The experimenter said, "I have some cards here with dots on them. I'm going to cover them up with my hand and then show you a card for just a second. I want you to tell me how many dots you saw on the card." The experimenter then withdrew her hand and exposed the card to the child for a brief interval (as brief as possible). This allowed the child time to see the card without allowing time to count the dots. The cards were presented to the child in random order and the child's subitising level was taken as the highest number on which no mistakes were made on four out of five trials, providing that no mistakes had been made with lower numbers. The children were continually encouraged but were not told whether their answers were correct or not.

Each child was then presented with tasks adapted from Bryant (1974) and requiring for their correct solution the application of the principle of invariance. The tasks were at two levels, subitising and counting, with four trials at each level. The eight trials were presented in random order. The procedure differed for control and experimental groups.

Control group: At the subitising level, the child was shown two rows of counters as in Figure 1a. The more numerous row contained a number of counters equivalent to the child's subitising score. The other row contained one fewer. The child was asked to say how many counters there were in one row, then to say how many there were in the other row, and then to say which row had the more counters in it. The experimenter then spread out the counters in the lesser row (see Figure 1b) and again asked which row had more counters in it. At the counting level, the procedure was the same (see Figure 1c and 1d) except that the child was told specifically to count (in the first comparison only, i.e., Figure 1c) and not simply asked to say how many as in the first task (Figure 1a).

FIGURE 1
TRANSFORMATIONS CARRIED OUT AT SUBITISING AND COUNTING LEVELS



Experimental group: The procedure was exactly the same as that for the control group except that pairs of counters were joined by a length of thick embroidery thread (see Figure 1). The counters were bevelled and the thread wound round them. This allowed for varying lengths of thread to be unwound so that the straight line of counters could be maintained when the counters were spread out.

RESULTS

For each trial given, the child received one point if he maintained his correct initial judgment and no point if he failed to do so. The maximum score at each level is thus 4. Table 1 shows the number in each group who reached the criterion of maximal performance (4 points) and the groups' mean scores over four trials in each condition. The difference between numbers passing all four trials in the control group and in the experimental group is very highly significant at both subitising and counting levels ($\chi^2 = 8.69$ and 13.57 ; $df = 1$; $0.01 > P > 0.001$ and $P < 0.001$ respectively). An analysis of variance of raw scores again shows a very highly significant difference between experimental and control groups and a difference significant at the 0.05 level between subitising and counting conditions. The interaction was not significant. The summary of this analysis is shown in Table 2.

TABLE 1
MEAN SCORES ON CONSERVATION TASKS AND NUMBERS OF CHILDREN WITH ALL-CORRECT RESPONSES

	Subitising		Counting	
	Control	Experimental	Control	Experimental
Mean score (max. 4)	1.33	3.73	0.60	3.33
Number of children passing all 4 tests	4	13	1	12
N	15	15	15	15

TABLE 2
ANALYSIS OF VARIANCE OF SCORES ON THE FOUR TASKS

Source	ss	df	ms	F	P
Total	200.6	59			
Between	179.6	29			
Experimental/control	106.6	1	106.60	40.89	<0.001
Error between	73	28	2.61		
Within	21	30			
Subitising/counting	3.27	1	3.27	5.51	<0.05
Interaction	1.13	1	1.13	1.91	NS
Error within	16.6	28	0.59		

DISCUSSION

There are two main findings. The first is that performance on these tasks is greatly facilitated by provoking a one-to-one correspondence between items in the two rows. This could be interpreted as support for Bryant's (1974) conflict hypothesis. Where the one-to-one correspondence cue is not immediately obvious (Figures 1b and 1d), conflict arises between previous judgments (Figures 1a and 1c) and the present array (Figures 1b and 1d). The more obvious cue is that of length and the child uses this to make his judgment. When the conflict is removed by provoking one-to-one correspondence the child is able to apply the principle of invariance very effectively.

The second finding confirms Young and McPherson's (1976) support for the Klahr and Wallace (1973) view that children first discover the conservation rules underlying invariance judgments for small numbers within their subitising range. The finding is also consistent with the Klahr and Wallace hypothesis that the conservation rules are discovered initially from application of the subitising operator and are later mapped onto the counting operator.

Both these findings have interesting practical implications. The first suggests a method for investigating conservation training; the second suggests that it may be necessary to teach number rules for small numbers before the logic of number can be apprehended. If this latter point is validated, it could mean that quite fundamental changes would be required in many modern approaches to the early teaching of number.

ACKNOWLEDGMENT.—The authors would like to thank the staff and pupils of King Street Primary School, Aberdeen, for their help.

REFERENCES

- BRYANT, P. E. (1974). *Perception and Understanding in Young Children*. London: Methuen.
 KLAHR, D., and WALLACE, J. G. (1973). The role of quantification operators in the development of conservation of quantity. *Cogn. Psychol.*, 4, 301-327.
 PIAGET, J. (1952). *The Child's Conception of Number*. London: Routledge and Kegan Paul.
 PIAGET, J. (1968). Quantification, conservation, and nativism. *Science*, 162, 976-979.
 YOUNG, A. W., and MCPHERSON, J. (1976). Ways of making number judgements and children's understanding of quantity relations. *Br. J. educ. Psychol.*, 46, 328-332.

(Manuscript received 3rd August, 1979)

Journal of Occupational Psychology

An international journal of research into people at work. Published quarterly, covering industrial, organizational, engineering, vocational and personnel psychology, as well as behavioural aspects of industrial relations and human factors. Innovative or interdisciplinary approaches with a psychological emphasis are particularly welcome.

Contents of Volume 53, Part 1, 1980

- N.A. Rosen, N.J. Georgiades & G. McDonald An empirical test of a leadership contingency model for teaching behavioural science concepts to managers
 D. Jacobson & M. Eran Expectancy theory components and non-expectancy moderators as predictors of physicians' preference for retirement
 R.A. Snyder Cross-lagged correlation analysis and the Zeltgeist
 J.A. Athanassou, P. Hall, G.A. Fox & J. Jenkins Classification of vocational interest factors on the Brook Reaction Test
 J. Cook & T. Wall New work attitude measures of trust, organizational commitment and personal need non-fulfilment
 A.J. Boyle 'Found experiments' in accident research
 R.J. Forbes & P. Jackson Non-verbal behaviour and the outcome of selection interviews
 C. Brotherton Paradigms of selection validation: Some comments in the light of British Equal Opportunities legislation
 A. Keenan & A.I. Wedderburn Putting the boot on the other foot: Candidates' descriptions of interviewers
 Book reviews

Volume 53 (1980) £25.00 (US \$55.00) Single parts cost £8.00 (US \$17.00)

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Hertfordshire SG6 1HN, UK

BOOK REVIEWS

FREEMAN, J. (1979). *Gifted Children: Their Identification and Development in a Social Context*. Lancaster: MTP Press, pp. xiii + 294, c. £6.95.

Although interest in gifted children and their education has increased considerably during the last decade there have only been a few research projects in this country concerned with them. The one described in this book by Joan Freeman, its director, known as the Gulbenkian Research Project on Gifted Children was funded by that Foundation and the initial selection of children in 1974 was made possible by the National Association for Gifted Children who allowed access to their accumulated records.

About half the book is devoted to a consideration of the research, the remainder providing a wide-ranging and well documented general introduction to the social development and learning of children with particular reference to those who are gifted, including their identification and a chapter about meanings of giftedness.

In the research project the 70 children who comprised the target group went to school in the north-west of England and each had a parent who had belonged to the NAGC; half had joined because of problems or possible problems, another third thought their child was 'different', a judgment that was justified as 33 of them had IQs greater than 150. The age range of the children was from 5 to 14 years and 84 per cent of them were below 11 years of age, so the study is largely about young children, and one deduces from the text that there were more boys than girls.

Joan Freeman matched these children, using Raven's matrices, for ability, and for age and sex with another child in the same school class. However, when these two groups were tested on the Stanford-Binet there was a highly significant difference between their mean IQs of 147 and 133 respectively. A second control group was chosen at random, again from the same class, and matched for age and sex, the mean IQ was 119; they were not a normal group as the author occasionally calls them. Thirty-seven per cent of the children were in independent schools and "... the parental population was of such an overall high social level that only six of the 61 schools were described by their headteachers as having generally working class populations". Indeed, the teachers were probably fairly used to teaching gifted children, for in addition to the 47 in the target group with IQs greater than 141 there were 26 in the first control group and 9 in the second; 71 per cent of all the children had at least one parent who was a teacher.

The parents of all the children were interviewed at home and a questionnaire completed about themselves and their children. The teachers completed a Bristol Social Adjustment Guide in respect of each child and the headteachers were interviewed.

The author discovered that the parents of the target children "... put very much stronger pressure on their children to achieve than either of the control groups" and 22 per cent thought their behaviour was very difficult compared with 3 and 7 per cent in the control groups; the teachers thought they were much more hostile and hyperactive. None of the children was unhappy at school; 80 per cent of the target group said they were very happy compared with 94 per cent and 84 per cent of the other two groups.

The main aim of the author "... was to find out if measured intelligence alone could be found to have behavioural effects, irrespective of parents' beliefs". To do so, she reallocated the 82 children with IQs greater than 140 from the three groups to form a high IQ group with a mean IQ of 155 and the remaining 128 to a moderate IQ group with a mean IQ of 120. Through doing so she was able to show from parents' descriptions that behaviour problems were noticeably absent in both groups and was able to present "... the most dramatic finding of the whole research. On the teachers' reports of the child's behaviour in school, *there was absolutely no difference between the High IQ and Moderate IQ groups* ... It cannot now be said, on the basis of these research findings, that intellectually gifted children are unhappy and show troubled behaviour just because of their gifts."

However, it is questionable whether many of the other findings would be applicable to gifted children who did not come from such a narrow range of the social spectrum and go to such unrepresentative schools.

J. WILKS.

GALLOWAY, D. M., and GOODWIN, C. (1979). *Educating Slow-Learning and Maladjusted Children: Integration or Segregation?* London and New York: Longman, pp. xii + 163, p. £3.95.

The recommendations of the Warnock Committee on Special Educational Needs have led to widespread discussion among teachers, advisers and psychologists: how feasible is it in practice to integrate children with handicaps into ordinary schools, what additional resources will be necessary, how profoundly will attitudes have to change, will there still be a need for some special schools? Galloway and Goodwin have made a timely contribution to this debate, concentrating on the largest category of 'handicapped' children—the educationally sub-normal and the maladjusted. They stress the non-medical nature of these difficulties and the important contribution of schools in determining the degree of handicap.

They begin with a historical over-view, drawing attention to the *ad hoc* development of special education as well as the almost total absence of evaluation studies. They make a strong case for a continuum of provision with an emphasis on integration whenever possible; special units or schools should be seen as a short-term treatment provision rather than as a protective alternative form of education. New developments in Scandinavia are described where schools as well as classes are significantly smaller than in England. Examples of successful integration in English schools are discussed and the need for early intervention such as infant school nurture groups is stressed.

The authors are at their best when they discuss the practical issues of classroom management, the essential extra resources needed, the need to adapt the curriculum (and worksheets) for the slow learner. The logical arguments with which the book begins are less satisfactory. The authors frequently shift their categories. Sometimes they refer to the small group of children for whom ESN schools were in fact planned: those children with IQs and attainments at least two standard deviations below the mean. At other times they are referring to the larger 10-12 per cent of dull children, most of whom remain in ordinary schools. The fact that these two groups overlap is not to say they are identical. This confusion leads them to set up a straw man—a psychologist who selects only on the basis of IQ. Burt, many years ago, said that a test was the beginning not the end of an assessment, and emphasised the multiple causality of backwardness. Chazan, Moore, Williams and Wright state quite clearly, "Educational sub-normality is a descriptive category rather than a diagnosis . . . the ascertainment of a child as ESN or slow learning is no more than the first stage of a process of enquiry to determine the reasons for the learning problem." The same can be said of maladjustment and bedevils all attempts at research in this field. These labels are as non-specific as having a temperature. A medical analogy makes this clear: most children with a temperature recover spontaneously; only a few are seen by a GP and of these even fewer are admitted to hospital. Some die, some recover and some are discharged with a chronic condition. Research to assess the outcome of treating hospitalised children with temperatures would not be very impressive, yet no-one would argue that because a majority remain at home hospitals are not needed or that there is no difference between those admitted and those who are not. Educational psychologists, like GPs, make clinical judgments and recommendations based on the resources available both in the school and the family. These will vary in time and place so that the populations of special schools will of necessity be heterogeneous. For each child the reasons leading to a placement and the aims of that placement will differ, making research based on diagnostic labels impossible and very easy to criticise.

Galloway and Goodwin are right, however, in drawing attention to the limited aims of some special schools and in urging education authorities and psychologists to put more resources and emphasis into helping ordinary schools with the education (not just the management) of their difficult and slow children. They would like to see the best aspects of special education more generally available. Although they recognise the difficulties of changing attitudes and the cost of extra resources, their optimism remains. Their book deserves to be widely read; it will inform staffroom discussions and heighten awareness of the complex adjustments, both educational and social, that schools will have to make if more than lip service is to be paid to the Warnock recommendations.

EVA HOLMES.

KAUFMAN, A. S. (1979). *Intelligent Testing with the WISC-R*. New York: Wiley, pp. viii+268, c. £12.25.

This book is addressed primarily to psychologists who may routinely or occasionally administer the WISC-R in the course of their professional work. The author sets himself two tasks. First, at a time when the mere mention of intelligence-testing is likely to evoke uncompromising and distinctly polarised responses, Kaufman states that he is attempting to "take a realistic look at the limitations of intelligence tests and to show that these flaws are not debilitating". Second, and this comprises the major part of the book, he aims "to integrate the theoretical with the practical", more specifically "by applying psychological theory and research to WISC-R interpretation, a systematic and step-by-step method of profile attack is advanced".

The first chapter deals with the advantages and disadvantages of intelligence tests. Kaufman's argument is that the practising psychologist must take a 'middle of the road' approach to testing, partly by reason of his statutory obligations. The professional should put intelligence testing in its proper place within a scheme of pluralistic and multifactored assessment and should be sufficiently aware of the shortcomings of, in particular, the WISC-R in order to be more incisive in the interpretation of an individual's test scores. This contribution to the philosophical, political and social controversies about the concepts of IQ and intelligence is slight, but many practitioners will appreciate his understanding of their problems and his plea for compromise. Perhaps his US readers will value his arguments even more: the frequent reference to legislation and litigation in his argument in favour of the use of the WISC-R may serve to reduce the confusion and wariness some American child psychologists must experience when deciding upon assessment procedure.

The remaining chapters of the book are a response to the second task which the author set himself. A strategy for constructing successively more refined hypotheses which, with the help of other forms of assessment, might lead to an accurate diagnosis of a child's difficulties is outlined. This is not a new idea; almost a decade ago Wedell, in this country, described a similar sequential strategy. However, what is new, and what is impressive, is the extent of the literature search undertaken by Kaufman and the way in which he presents information from this literature, from his own work, and from sample cases, in a clear, concise and practically useful format.

Chapter 2 concentrates upon differing interpretations of Full Scale IQ and Verbal-Performance discrepancies. Their relationship is examined in terms of factor analytic studies, Cattell and Horn's theory of crystallised and fluid intelligence and Guilford's operational model of intelligence. He assesses the possible effects of factors as diverse as time pressure and bilingualism, re-testing and cognitive style and provides comprehensive case reports to illustrate the points he makes.

Kaufman's own factor analytic studies have revealed a 'freedom from distractibility' factor present in the WISC-R along with Verbal Comprehension and Perceptual Organisation factors. Possible interpretations of the third factor are the subject of a complete chapter and observations about its relationship with test anxiety and mental retardation seem well supported. The failure of other factor analytic studies to produce such a factor, however, may lead to some reservation about its interpretation.

Further chapters attempt to identify the abilities measured by the WISC-R and the extent to which such abilities are general or specific; provide detailed analyses of different subtest score patterns within the verbal and performance scales; and finally a specific six-stage strategy for interpreting subtest fluctuations and scatter is proposed. Again, illustrative case studies are provided together with recommendations for action.

The argument for the continued use of intelligence tests is not totally convincing and will certainly not be accepted by all. The book, despite its rather 'slick' title, makes a reasoned and substantial case for the continued use of the WISC-R as an aid to diagnosis. For those who, like the reviewer, feel that such a technique can still contribute to our assessment of individual children, it is essential reading and should attract many individual sales as a source of ideas about interpretation that might otherwise have remained unconsidered.

IAN BERRY.

KLINE, P. (1979). *Psychometrics and Psychology*. London: Academic Press, pp. v+381, c. £19.00.

Paul Kline's book reviews the findings of psychometric research with the joint aims of demonstrating the contribution of psychometrics to psychological knowledge and indicating its relationship to psychological theory. The book begins by defining psychometrics as that area of psychology concerned with testing, and proceeds to discuss in detail the usual test-related issues of validity, reliability and standardisation, also outlining 'the art' (item writing) and 'the science' (the selection of items, and the demonstration of reliability and validity) of test construction.

Chapter 2 deals with the statistical procedures adopted by psychometrists, providing a very brief description of the processes involved in a variety of multivariate analysis techniques, followed by a fuller account of factor analysis and different factor analytic designs. Kline's discussion of factor analysis, however, presupposes some familiarity with the technique and since this is the main statistical tool of the psychometrist, a more complete statistical, or at least conceptual, account would perhaps have been helpful to the student reader. This is followed by a chapter describing the psychological tests themselves which are subdivided into six categories: intelligence, aptitude, attainment, personality, projective and objective.

The succeeding three chapters set out the substantive findings resulting from the application of these tests: first, in the area of human abilities, in which Kline argues strongly against the popularised view that intelligence is what intelligence tests measure, and quotes evidence strongly supportive of the construct validity of many psychometric factors; then chapters 5 and 6 take a similar format for empirical research in the areas of human temperament and human dynamics.

Chapter 7 relates psychometric findings to theories of learning, more specifically to the theories of Cattell and Eysenck. The following three chapters deal with the applications of psychometric findings to the fields of education, clinical psychology and cross-cultural psychology. Here Kline presents highly speculative and less convincing arguments of the potential of psychometric research. Psychometric tests, for example, are viewed as powerful diagnostic instruments for the educational and clinical psychologist, and are expected to lead to the successful matching of people to jobs in vocational guidance. In these controversial chapters Kline makes some ambitious claims for psychometrics and appears at times to have lost sight of his earlier admissions of the imperfect reliability and validity of psychological tests.

The book concludes with some speculations on the contribution of psychometrics to a theory of human behaviour, and the suggestion that psychologists, in their construction of theory, become more cognisant of psychometric research.

Overall, Kline presents a well-organised and extensive review of psychometric techniques and findings, and the book, though expensive, could provide a useful text for psychology students. However, in attempting to relate psychometrics to psychology in general there are two important weaknesses. First, psychometrics is defined in terms of a set of methods and statistical processes, and Kline advocates that theory be built upon the findings which result from their use. Yet if Kline is to overcome the criticism of basing theory upon, at least initially, blind empiricism, a much closer examination is required of the theory and assumptions which give rise to psychometric research itself, together with a rationale for why such a methodology should be deemed fruitful. Second, Kline is somewhat intolerant of other methods of research, assuming that psychometrics, and psychometrics alone, will lead psychology to advances emulating those of the natural sciences. As a result, the appropriateness of different methods in different research contexts and the potential complementarity of different approaches are not fully explored. At times, it seems that Kline provides us with less of an attempt to link psychometric findings to different areas of psychology as to reduce psychology to psychometrics.

JAMES CALDERHEAD.

LIGHT, P. (1979). *The Development of Social Sensitivity*. Cambridge: Cambridge University Press, pp. 124, £7.95.

Piagetian theory has been questioned of late by a number of writers. Its exclusive concern with cognitive aspects, its hard and fast exposition of developmental stages, and the

nature of some of the specific concepts on which the theory rests are all increasingly being challenged. Among the concepts which have been most vulnerable to further empirical scrutiny is egocentrism: the unqualified assertion by Piaget that children, before they become capable of operational thinking around the age of 7 years, are so embedded in their own point of view that they are unaware of other people's different orientations and feelings. In Piagetian theory this concept assumes in many respects a central role; the failure to find its behavioural manifestations in the spatial, verbal and social areas in an absolute form is therefore of serious concern.

This book reports a further series of studies in which egocentrism is examined. It does so from the point of view of the child's role-taking ability, i.e., the extent to which he can make inferences concerning another person's feelings, expectations, attributes and likely reactions. A variety of measures were used to test this ability and applied to four-year-olds. The scores so obtained showed significant intercorrelations, even after partialling out IQ, and a composite score was therefore derived which could be related to various possible antecedent and consequent conditions. The emphasis in this report is thus very much on individual differences—another aspect to which Piagetian research has given little attention. In particular, various characteristics of the children and their families were investigated in order to search for concomitants of the marked differences in role-taking skill found in the sample. A number of interesting and thought-provoking findings emerge, though considering that the research derives from the Richards and Dunn longitudinal study in Cambridge, with its rich data base for each family extending over several years, a rather more liberal use of that knowledge could have been expected. Certainly measures derived from answers to single interview questions are not likely to provide any degree of impressive validity. In addition, the correlation of role-taking skill at age four with various single indices in infancy presents a research strategy that is now no longer considered appropriate in longitudinal investigations. It suggests a once and for all early determinant the influence of which one is bound to pick up years later. As the present interest in transactional models of development suggests, the child's psychological growth is too complex to make such an approach even worthwhile attempting.

Nevertheless, Light does a service in placing cognitive development squarely within the social context in which it surely belongs. He presents his data modestly and clearly, and has no hesitation in pointing to gaps and deficiencies. It is the kind of research that one is perhaps more used to reading about in journal articles than in book form: whether a report so specialised and (given its brevity) so expensive will find a sufficient readership to make it financially worthwhile is a problem that the publishers, hopefully, have come to terms with.

H. R. SCHAFFER.

LUNZER, E. A., and GARDNER, K. (Eds.) (1979). *The Effective Use of Reading*. London: Heinemann Educational Books for the Schools Council, pp. 337, £4.50.

From 1973 to 1976 the Schools Council supported a major study of the reading of 10 to 15-year-old pupils of average and above-average reading attainment. This book provides a coherent and stimulating account of the research team's work.

The content falls into three main sections. The first (chapters 2-4) begins with a theoretical analysis of the relationship between spoken and written communication and the problems involved in comprehending textual material. Against this background, an empirical study of the characteristics of reading comprehension based on a sample of 257 fourth year junior school pupils is described. On the basis of their findings and a review of the research literature, the authors conclude that reading comprehension comprises "a single aptitude or skill, one which cannot usefully be broken down" (p. 63). Individual differences in reading comprehension are a function of the pupil's "... ability and willingness to reflect on whatever it is he is reading" (p. 64). A study of the readability of textual material used in various aspects of the curriculum and of the utility of readability measures leads to two main points. The first is that the texts read by pupils when under a teacher's supervision in class

tended to be *less* difficult than texts studied unaided. Secondly, readability formulae are of value to the teacher when used with an appreciation of the limitations of such procedures.

The second section (chapters 5 to 7) describes the incidence and context of classroom reading, reading done in connection with homework and reading carried out in relation to topic work. Marked differences between the pattern of reading done in the top class of the primary school and the first year of secondary school reading are specified. Of considerable importance is the finding that school reading is characterised by 'short bursts' of from 1 second to 15 seconds in any one minute in about 50 per cent of all reading across all subjects. Such intermittent reading is seen as incompatible with the reflective thinking that the authors consider the essence of reading comprehension. More continuous reading was required in homework tasks but the authors question whether pupils are taught how to do this effectively, although it is claimed that it can be done. Turning to topic work, whilst this is seen as potentially conducive to reflection on what is read, doubts concerning current practice are expressed. The promotion of effective reading requires more of the teacher than lighting the blue touch papers of pupils' enthusiasms and retiring to a safe distance. Gardner comments: "The reference book was treated more as a treasure trove of sentences and paragraphs which could be stolen and marketed again in another setting" (p. 305). This observation rings bells and is a phenomenon found not only at the 10 to 15-year-old age level.

The last section reports on school-based practices having considerable promise in promoting the more effective use of reading. In chapter 8, the efficacy of the SRA Reading Laboratories in developing particular aspects of pupils' reading is found to be convincingly displayed. "If this conclusion is to be attacked, then evidence must be produced to show where our experiment was defective" (p. 307). Challenging words! The apparent contradiction between the advocacy of the unitary nature of reading comprehension and the value of a skill-based system such as that embodied in SRA Reading Laboratories is discussed and resolved to the satisfaction of the authors. The Reading Laboratory material enhances the willingness of pupils to reflect on what has been read. Having established the value of the SRA materials, the researchers found that their hard-won knowledge was met in many instances by scepticism and even antagonism. One is reminded of a similar reaction found during a study of the value of a structured programme of language development activities at the nursery level (Quigley, 1971). Pilot work on five related approaches to the improvement of reading through group discussion is described in chapter 9. The promise of these approaches is considerable. The Schools Council has, in my opinion, shown wisdom in funding further work needed to develop, disseminate and evaluate the use of such strategies in various subjects.

This book is one that merits the attention of those interested in the nature of reading comprehension and the development of approaches that will enable teachers, irrespective of their subject specialism, help their pupils read more effectively.

The publication is likely to produce a considerable amount of constructive academic dialogue, although doubtless both heat and light will be generated in the event. What might be regarded as some of reading's current sacred cows suffer a certain amount of blood-letting, if not slaughter.

In a restricted review it is not possible to do justice to detailed technical aspects of this major report. However, on the key issue of the nature of reading comprehension, some comments can be made. Proponents of the skills analysis approach to improving reading will question whether reading comprehension is as unitary as the Nottingham team believe. Are those in the opposing camp totally incorrect? The Nottingham team's data are based on the 10 to 11-year-old age group and then, through reference to the research literature, the unitary nature of reading comprehension is generalised to a wider age range. It can be argued that children's cognitive abilities continue to differentiate during the period from 10-15 years. Further, no mention is made in the report of an important empirical study of listening and reading comprehension carried out under the aegis of the Australian Council for Research in Education (Clarke, 1973). In the context of the evidence available, unanimous agreement with Lunzer's interpretation by all specialists in this field is highly unlikely. Lunzer does accept the possibility that the sub-skills hypothesis might be vindicated, but is dubious as to the potential utility of such a finding (p. 65).

On reflection, it seems possible that both the comprehension sub-skills approach and the integrated-reflective approach to reading comprehension may contain points of value in

describing the nature of reading comprehension and facilitating its development. Clarke's work points towards such a synthesis.

In his 1974 Presidential Address to the British Psychological Society, Professor B. M. Foss spoke on the value of taking sides in the advancement of knowledge. The nature of reading comprehension and the means by which it might best be fostered are crucial theoretical and practical issues. The Nottingham team is to be congratulated on presenting its analysis of these issues. Their position is clear. The response of those with reservations concerning the various studies and interpretations reported should further our conceptualisation and control of the development of reading comprehension—eventually.

PETER D. PUMFREY.

REFERENCES

- CLARKE, M. L. (1972-73). *Hierarchical Structure of Comprehension Skills, Vols. 1 and 2*. Victoria: Australian Council for Educational Research.
 QUIGLEY, H. (1971). Nursery teachers' reaction to the Peabody Language Development Kit. *Br. J. educ. Psychol.*, 41, 155-162.

MITTLER, P. (Ed.) (1978). *Research to Practice in Mental Retardation. Volumes I, II and III*. London: MTP Press, £14.95 each.

These three volumes contain 160 papers given at the 4th Congress of the International Association for the Scientific Study of Mental Deficiency held in 1976. A title list of 300 papers not selected for inclusion indicates the demanding task undertaken by Professor Mittler in editing these volumes. The aim of the Congress was to relate research findings to practice, to lessen the wide gap between research workers and practitioners; during informal discussions at the Congress such an interchange was no doubt possible. Nevertheless volumes II and III in particular will be more valued by research workers than by field workers.

Volume I is concerned with Care and Intervention. It is the most in touch with practical issues such as developments in residential and community care, and attitudes to the retarded and their families. It has an excellent section on early intervention in which Stedman provides a comprehensive and sensitive review of the problems involved. There is an account of the important Milwaukee project and a follow-up of Gray's work in Nashville with mothers of young children in poor families. This section, as well as the introductory paper by Professor Clarke, draws attention to an important distinction often ignored by others—that between the causes of severe and mild retardation. The prevention of severe retardation is primarily a bio-medical problem whereas mild retardation is now recognised to be associated with complex social and environmental factors.

This is highlighted by volume III—devoted to Bio-medical Aspects. This volume is of limited interest to psychologists because it is mainly concerned with very technical bio-medical papers on such conditions as phenylketonuria, Down's syndrome and metabolic disorders. It draws attention to recent research on congenital malformations and new chromosomal anomalies. It underlines the need for ante-natal screening and specialised diagnostic facilities. There is an important section on malnutrition. A paper by Scheiner *et al.* indicates that even children with a congenital infection (CMV) associated with increased risk of hearing loss and intellectual impairment more often found in young families of low socio-economic background, were less affected when they had some nursery experience. Another paper by Porges and Humphrey demonstrates some differences in physiological responses, associated with sustained attention, between normal and retarded children.

Volume II, Education and Training, is disappointing. The sub-sections reflect research interests: cognition and learning, behavioural intervention, epidemiology. A practitioner looking for papers about children, or adults, or severe or moderate subnormality, or parents, would have difficulty in finding what he needed. Two-thirds of all the papers in this collection are American; in volume II in particular these are often characterised by a pretentious use of language which very effectively distances the reader from the reality of the mentally retarded people in long-stay institutions who are being talked about: instructional delivery

systems, behavioural deficits, salience of stimuli, mutualistic interaction, cognitive modifiability, the 'learner as data source' orientation. Very rarely does anyone write about individuals, their needs or relationships, let alone their feelings. Careful studies are reported on how to teach the 'target population' to read a few words, to memorise three rather than two digits, to discriminate shapes. Often these studies raise points of theoretical interest but their relevance to the life of the 'subject' is not clear. Variables that are difficult to measure are almost always ignored: relationships with staff, attachment to peers, contact with relatives. This is sometimes a matter of satisfaction: "Very little need be known about the specific individual, except that he performs certain behaviour producing a profile" (p. 154) or: "Parental visitation is not necessary for maintaining desired resident behaviour, but may actually be detrimental". A familiar viewpoint expressed by nurses in children's wards in the 1940s! In contrast to this type of research report, the accounts of practice in Sweden, Israel and Scotland are refreshingly jargon-free and down to earth; they are helpful and informative to those involved in planning services. Thoughtful reviews on general issues such as 'mainstreaming' were also welcome.

There is no doubt that the scientific study of mental subnormality has made a very positive impact on the morale and quality of life in long-stay institutions. Changes in expectations, staff, organisation, finance and rehabilitation policies have come about. A gap—perhaps inevitably—remains between academic research interested in processes such as learning and perception, and the practical surveys and projects being tried out in a particular hospital or area which need to be reported to other practitioners. These volumes have not been wholly successful in attempting to bridge this gap.

EVA HOLMES.

OLWEUS, D. (1978). *Aggression in the Schools, Bullies and Whipping Boys*. Chichester: John Wiley, pp. xiii + 218, c. £9.95.

Any book which sheds more light upon the problem of the bully and the bullied in schools is welcome. Dan Olweus's efforts in this field are considerable. His work on aggression in school provides a major contribution to our knowledge and understanding of the psychological and sociological factors involved in the whipping boy/bully problem.

He has presented in his comprehensive book an outline of his research findings on the incidence of whipping boy and bully problems within the Norwegian school situation. With considerable attention to detail he outlines the methods of his studies which have been carried out in the Stockholm area and he allows the reader an opportunity to critically appraise his methods and his results.

This is certainly not a book in which the author has stinted on detail. It is to his credit that at the end of each chapter he has re-presented the material in a concise and readily assimilated fashion; this will undoubtedly help those readers who are less familiar with the intricacies of psychometric, psychological and sociological investigations, and allows the reader to tune in to the main points of his study.

It comes as no surprise to find that the whipping boy/bully problems seen in school seem to be largely bound up with the character and nature of the interpersonal relationships which exist between the pupils in the classroom. I was, however, personally surprised to find that the research evidence showed that there was little relationship between the problem of bullying and the size of classes and schools. To me this seemed particularly strange when I compared it against a background of my own clinical experience in trying to help teachers manage similar difficult situations in large comprehensive schools. We readily assume that we can help the child and the school situation by placing the vulnerable child in a smaller school and particularly in smaller classes where there is higher staff/pupil ratio. The research studies outlined in the book indicate that physical strength and aggression present in an exaggerated form in school children are more frequently found in whipping boys rather than their ordinary counterparts within school. One of the most depressing features of Dan Olweus's study has been highlighted in his research findings during a three-year follow-up period of those boys who had problems with their aggression, either as bullies or being bullied. They showed very little change during this period and this was despite a

considerable amount of environmental variation and manipulation carried out by teachers concerned in their attempts to alleviate the problems in the child under their care.

I was disappointed to find that only a small proportion of the book was devoted to suggestions about how we can help the bully and bullied within the school environment. Hopefully, in a subsequent volume, Dan Olweus will develop his views on this, set against a background of quite extensive research in this field. I felt that this book provides a useful contribution to the understanding of behavioural problems in schools and that it is more likely to be read by those academically inclined in education rather than the teacher who is looking for specific practical experience in the management of children within the classroom.

MARTYN J. GAY.

PERROTT, E. (1977). *Microteaching in Higher Education: Research, Development and Practice*. Guildford: S.R.H.E. Monographs, pp. 107, p. £3.75.

Since Professor Perrott is the most eminent British academic to have taken a sustained interest in microteaching, this monograph, in which she gives an account of much of her work in the area between 1969 and 1977, is of considerable interest.

The title is somewhat misleading, since the monograph is concerned with the use of microteaching in the professional education of teachers in primary and secondary schools. The major focus is on an evaluation of a 'minicourse', adapted for British conditions by Perrott and her colleagues from an original version developed and tested at the Far West Laboratory, San Francisco, and used for in-service professional education. There are, however, two preliminary chapters.

The first chapter offers a generalised description of microteaching, uncompromisingly behaviourist in its orientation, and a brief review of relevant research. It would be unreasonable to expect, in the few pages devoted to it, a balanced or searching research review. The reader may be disappointed, however, by the little account taken of the large number of research publications since 1971, and by an apparent tendency for the research to be used to justify the minicourse design: for example, emphasis is put on the lack of clear evidence indicating the value of supervisory feedback (not provided in the minicourse), but no mention is made of the comparable lack of evidence showing the value of the practice phase of microteaching (included in the minicourse).

The second chapter gives an account of the use of microteaching in the early years of Stirling University's concurrent teacher education programme. Since Professor Perrott was the first Professor of Education at Stirling, this account gives the perspective of one of the most important participants during this innovative period and is therefore of historical interest. From the perspective of this reviewer, the picture which emerges looks too much like a blueprint which was automatically translated into practice, with no sense being given of the divergent interpretations, the conflicts and the constant changes which occurred; but then, the value of the account lies in the fact that it *does* give a distinctive perspective.

The meat of the monograph, however, is in the report of an investigation into the effects of the adapted version of the 'Effective Questioning' minicourse. In relation to behavioural measures of the effects of the course, "the main field study of the British adaptation followed the outlines of the American main field study (Borg *et al.*, 1970), and produced results that are remarkably similar, even to the point of obtaining the largest and most stable differences for the same measures" (p. 47). In a study which looks in most respects to have been technically careful and competent, teachers were videotaped during 20-minute discussion lessons in their normal teaching environments before and after taking the course, and again four months later. On eight of the 14 criteria relating to the objectives of the course, significant gains appeared to have been made and maintained, and these observed differences correspond quite closely with teachers' own reports on ways in which their teaching had been influenced by the course.

A problem in all studies of this kind is that of knowing whether teachers' *normal* teaching has been influenced by the course they have taken, a problem which Perrott correctly relates to the apparently inconsistent findings reported by researchers on the 'stability' of teachers'

behaviour. She is to be congratulated on having collected some evidence relevant to this problem; when teachers were asked to comment on the effects on their behaviour of being videotaped, their reports suggested that during the pre-course lesson they used the relevant skills less than they normally would have done, but that during the post-course lessons they used these skills considerably more than they normally would. It could therefore plausibly be suggested that the measured increases in their use of the skills were due solely to their increased confidence in front of the camera and to their clearer understanding of what the course evaluators were looking for, not at all to any generalised change in their teaching behaviour. Perrott's results thus highlight the ambiguity and inadequacy of virtually all research findings concerned with the effects of training on teachers' normal classroom behaviour. It seems unlikely that these deficiencies can be overcome within the framework of a conceptualisation of the regularities in teachers' activities in terms of the stability of counts of behaviours, or so long as courses are seen as being aimed at changing teachers' behaviours. A more fruitful way of approaching the issue may be to consider such courses as being aimed at influencing teachers' ways of judging how best to act in specific situations and at enabling them to act in ways they consider appropriate; the achievement of such course goals could not be assessed through the simple counting of behaviours.

A further chapter in the monograph goes on to explore teachers' attitudinal reactions to the minicourse. Many of the findings reported are useful in providing a profile evaluation of the specific course, but their more general interest is limited. Perhaps the most thought-provoking finding is that teachers tended to see themselves as having been influenced towards a less formal general style of teaching, although this was not one of the explicit purposes of the course: the relationship between planned and actual outcomes of courses can be complex. It is surprising that this is not mentioned in the final chapter's discussion of the virtues of a systems approach to teacher education.

Perrott does not discuss in this monograph the question raised by Applebee (1976), one of her collaborators in the work described, about the doubtful wisdom of spending large amounts of money to adapt American materials for use in Britain. Recent findings by Batten (1979) suggest that, without modification, Far West minicourses can be highly effective in the British context. The research findings reported in this monograph may therefore best be interpreted as demonstrating, together with those of Borg, Batten and others, the effectiveness of the very careful and highly structured Far West approach in at least sensitising teachers to potentially important aspects of their classroom teaching.

DONALD MCINTYRE.

REFERENCES

- APPLEBEE, A. N., (1976). Microteaching, component skills and the training of teachers: an evaluation of a research and development project. *Br. J. educ. Technol.*, 7 (2), 35-44.
 BATTEN, H. D. (1979). Factors influencing the effectiveness of microteaching in a teacher education programme. Unpublished Ph.D. Thesis, University of Stirling.
 BORG, W. R., KELLEY, M. L., and LANGER, P. (1970). *Minicourse 1: Effective Questioning*. Beverly Hills, California: Macmillan Educational Services.

REX, J., and TOMLINSON, S. (1979). *Colonial Immigrants in a British City. A Class Analysis*. London: Routledge and Kegan Paul, pp. xv+357, c. £10-50.

The authors claim that in this book they present 'a self-consciously reflective sociology of race relations in Birmingham' (p. 320). This is a complex work with a variety of themes running through it. Firstly, there is a good deal of discussion, *inter alia*, of the methodological problems of sociology in general, and of the pitfalls and hazards of studying race relations in particular. John Rex's concern with these theoretical difficulties is apparent, and, indeed, is revealed most clearly in his contribution of an appendix to the book entitled 'Sociological theory, typologies, value standpoints and research'. The authors eschew 'value neutrality'; nevertheless they strongly condemn "those, who mistake the process of myth-making for sociology and who would demand of us that as a mark of our good faith we assert to propositions which we know to be untrue" (p. 310). It is difficult to categorise

briefly the authors' outlook: perhaps 'marxist' might do, but they are eager to distinguish themselves from the more vulgar and dogmatic marxist.

Secondly, there is at the core of the book an account of a social survey commissioned and largely designed by the authors, but carried out by National Opinion Polls, on the immigrant and non-immigrant population of Handsworth—a multi-racial area in Birmingham. The results of this survey provide much information on the attitudes of immigrants and non-immigrants to the problems of employment, housing and education. In general it appears that immigrants are rather more happy with living in Handsworth and with their employment than the whites in the area, who are somewhat more inclined to blame alleged deficiencies on the presence of immigrants. Even the sample of 'immigrant' youth appears at least relatively content, and for all their well-publicised conflict with the police are prepared to leave law enforcement to them.

These findings are set within a discussion in which the overall view is that the sociological analysis of race relations must place such relations in a wider context than Britain alone. In this analysis, the authors suggest that the British empire, and the class relations which occurred within it, is of fundamental importance. This overview leads Rex and Tomlinson to find no grounds for optimism in the results of their survey. Indeed, they are strongly critical of those who take an optimistic view of race relations: for example, politicians and policy-orientated sociologists. They are almost as critical of liberal approaches to race relations as is Enoch Powell, and, indeed, one is tempted to conclude that they share his apocalyptic vision of the future of such relations in this country. There is much to ponder over in this book, but one might wonder whether this extraordinarily rich amalgam of theoretical discussion, policy suggestions, empirical enquiry, and methodological speculation is not a curiously muffled and confused contribution to the literature on race relations.

B. SHAW.

SIEGLER, R. S. (Ed.) (1978). *Children's Thinking: What Develops?* Hillsdale, New Jersey: L. Erlbaum, pp. xi+371, c. £13.30.

There is no doubt that much contemporary American research into children's intellectual development is of a very high standard, and this book contains a valuable sampling of the best. The editor has little to say about the individual papers, which he briefly summarises in a short preface, but he also contributes a chapter on the origins of children's scientific reasoning. There are three sections, on memory development, problem-solving and representational processes. Each section contains three or four chapters by researchers deservedly prominent in their respective fields, followed by a discussant's comments. Thus Flavell discusses a chapter by Brown and DeLoache on skills, plans and self-regulation, one by Case on intellectual development from a neo-Piagetian perspective and a chapter by Chi about knowledge structures and intellectual development. Greeno comments on papers about problem-solving, by Siegler, by Trabasso and others (on solving class-inclusion problems), Klahr (goal formation, planning and learning) and Gelman (counting in pre-schoolers). Finally, G. H. Bower discusses articles by Nelson, on the representation of knowledge about the everyday world, Gregg on spatial concepts, Kosslyn, on imagery and cognitive development, and by D. P. and H. A. Simon, on individual differences in solving physics problems.

Bower remarks of his role as a public discussant: "I have found it a most disagreeable job. I dislike the role of examining someone's prized writing, looking for possible flaws or things to take issue with or disagree with." He is disenchanted with the 'adversarial' view, according to which debate and criticism of scientific ideas, with an emphasis on finding the weakest spot in someone's arguments, is regarded as the best way to advance Science and Truth. Bower argues that such an approach "leads to devastating criticisms that are terribly destructive of the spirit, self-esteem and enthusiasm of the parties involved as well as the onlookers. First and foremost, scientists are human beings whose productive work requires sincere praise and encouragement from peers in order to maintain enthusiasm for their scientific pursuits." Hear hear!

All the contributors to this book deserve praise and encouragement. The chapters display a concern with the kinds of abilities that are important for the child's confrontations with the real world. This book will be valuable for psychologists who wish to deepen their understanding of the development in children of thinking and the abilities that depend upon thought.

MICHAEL J. A. HOWE.

SKEMP, R. R. (1979). *Intelligence, Learning and Action*. Chichester: John Wiley, pp. xv+324, c. £14.50.

Psychologists generally have recognised that the traditional notion of intelligence as an innate mental power, measurable by intelligence tests, is no longer adequate, and instead are beginning to analyse and experiment on the psychological processes involved in intelligent thinking and behaviour. Dr. Skemp's book is a good example of this new approach, and he provides a comprehensive survey of cognitive and affective processes underlying human growth and learning, from the simplest to the most complex. Somewhere intelligence itself disappears in the course of his arguments—perhaps this is intended—though he does go into great detail in analysing the more and the less intelligent types of learning, particularly in school. The nearest he gets to a definition is: "An advanced kind of learning, by which director systems are built, tested, up-dated, and improved in their internal organisation." (Director systems are internal controls which direct the individual's actions towards the attainment of his goals.) I am unhappy about identifying intelligence as a 'kind of learning', although agreeing with him that children born with appropriate genes, who obtain a stimulating upbringing, are better able to learn intelligently.

It is quite impossible, in a short review, to summarise the main points in Skemp's theory of thinking behaviour, and indeed I am not sure that I have understood him sufficiently to attempt this. The book is excellently organised and well written, but it is very tough going. I found it difficult to maintain attention to more than a few pages at a time. Dr. Skemp realises this complexity, and recommends that the book should be read twice, which would certainly help. Where he differs from other writers is that he makes minimal use of experimental evidence. His analysis is almost entirely observational and introspective, and yet I generally find it convincing. It is noteworthy that his bibliography amounts only to about 90 references, and many of these are from non-psychological sources. On the other hand, a glossary covers some 130 technical terms, which he has invented or adapted, to serve as concepts in his theory. One would indeed get lost without frequent reference to this glossary. While I have referred above to Skemp's contribution as a 'theory' (or model) of intelligent functioning, many fruitful applications to education occur in most of the chapters.

A useful way to categorise the book is to indicate some of the main sources of his views. Spearman's and subsequent factorial studies are briefly recognised. Piaget is more criticised than accepted, but Hebb and J. McV. Hunt are considered valuable. Miller, Galanter and Pribram's theory of 'plans' and 'totes' is a major source, but information theory (apart from a few concepts like 'noise') is ignored. Curiously, no reference is made to Ausubel, although Skemp's theory of school learning overlaps very considerably with his. All in all, this is an important book, which should be studied by all educational psychologists. But it would be too difficult for students, other than those working at Honours or Ph.D. levels.

PHILIP E. VERNON.

STONES, E. (1979). *Psychopedagogy: Psychological Theory and the Practice of Teaching*. London: Methuen, pp. xxii+490, c. £8.95, p. £4.95.

Professor Stones' new book is a radical venture in the use of psychology in teacher education. It is radical not so much in what is new as in what is dismissed.

Stones has imported the concept of psychopedagogy from Europe. Psychopedagogy is concerned with the application of psychological principles to the practice of teaching. This evidently precludes attention to the more general notion of education or even, it seems, to a consideration of the school as an institution. In Stones' view, if psychology has not got direct relevance, that is practical implications for the teacher in her role as instructor, then it is not psychopedagogy. Not surprisingly then, the book is light on theory and data and heavy on prescription.

Professor Stones' persistent theme is that the teacher should think systematically about her behaviour as instructor with a view to optimising that behaviour. As a framework in which to think he offers the 'objectives' approach. The teacher must be clear about her goals and about her task organisation. In this context Professor Stones introduces and uses the work of the Behaviourists. Additionally, Stones has developed a number of strategies for task and skill analysis and these are presented together with detailed examples of the strategies in use.

In the active phase of teaching the teacher must learn to monitor her own behaviour. In particular she must concentrate on examining the language she uses and on her role as a distributor of reinforcement. The psychology used to inform the inspection of classroom language arises from studies of language, thought and communication. Here Stones relies almost entirely on the work of Russian psychologists. His treatment of the teacher as reinforcement-manager is Skinnerian.

Finally Professor Stones discusses the evaluation of children's learning. Here again his emphasis is on what the teacher should be about. Attention is consequently focused on teacher designed, criterion referenced tests, their implementation and evaluation.

The book is clearly written and liberally illustrated with extracts from real lessons which the author uses as case study materials to make points, or to demonstrate the exercise of skills he seeks to communicate to his readers. The author's style also conveys his commitment to children and teachers. But Professor Stones is not only prescriptive at the level of particular teaching skills; he has a prescription for teaching in general. This is that teaching should be seen as a branch of experimental psychology in which the teacher engages in systematic alterations of her behaviour and carefully monitors the consequences. In the context of this overarching conception of teaching he examines and ruthlessly dismisses a large number of concepts and bodies of work. Intelligence, personality and cognitive style are all consigned to the educational rubbish bin. The book contains only two citations for Piaget. Stones' new field of psychopedagogy has little place for studies which separate development from learning.

Stones' book is avowedly a book for teachers in training. In this context his parsimony will no doubt be welcomed. But it is difficult to see how his experimental psychology of teaching will make progress as a science. Whilst his prescriptions may well produce improvements at the behavioural level, it seems that all problems at the conceptual level have been resolved. Stones gives no account of how his approach will add to the teacher's *comprehension* of what goes on in the name of teaching. It is interesting to note that most of Stones' prescriptions are at least 100 years old and that almost all of his case study materials are examples of 'how not to do it' because, evidently, it is very difficult to find exemplary teaching actions on record. This might have given Professor Stones pause for thought. What psychopedagogy will have to be about is the construction of a science of teaching which conceptualises real time practical constraints on teachers and treats them more seriously than is done here. To see psychopedagogy as less than this is a waste of a term.

CHARLES DESFORGES.

WALLER, G. T., and MACKINNON, G. E. (Eds.) (1979). *Reading Research: Advances in Theory and Practice, Volume I*. New York: Academic Press, pp. v+262, c. £12.00.

This book is the first in a series intended to bring together important contributions from various disciplines involved in studying the reading process and the assessment and alleviation of reading difficulties. Each volume will focus on a particular topic within these

fields and provide "systematic and substantive reviews and syntheses, both empirical and theoretical . . .". Ambitiously, the editors of this particular volume have chosen as its theme the question "What does a child have to know (or, be able to do) in order to learn to read?"

The first of the seven contributors considers the origins of the notion of reading-readiness. After a re-examination of the early work of Morphet and Washburne and of Dolch and Bloomster, the question "When can children learn to read?" is dismissed as unanswerable and the concept of reading readiness deemed redundant and even harmful. A case is made out for directing attention to extrinsic influences on learning. The chapter is well documented and argued.

The second contributor acknowledges that no agreed taxonomy of pre-reading skills exists but sees some advantages in building on the work of pioneers in this field. Four commonly used reading readiness tests are described, three normative and one a mastery test. The wide range of sub-skills purportedly assessed by eight commonly used reading readiness tests is presented in tabular form to illustrate the heterogeneity of skills sampled. Perhaps insufficient is made of the problems of interpretation that arise when looking at a profile of test scores. Insufficient emphasis is given to difficulties arising from sub-tests in different reading readiness tests having the same name but not necessarily tapping the same processes. Readiness tests, it is argued, "... should be constructed and used for instructional rather than predictive purposes". Other changes needed in readiness tests and based on recent research findings are specified.

A substantial, detailed and stimulating chapter on linguistic insights pertinent to our understanding of children at the threshold of reading acquisition follows. The relationship between 'word consciousness' and reading is discussed in the light of a series of experiments. Some intriguing bi-modal distributions of word-discrimination scores among pre-readers are reported and their significance discussed. The author is generally sympathetic to the notion that 'learning to read' facilitates the child's awareness of various aspects of the relationship between speech and print. Contrasting theoretical positions are discussed and some common ground identified culminating in a recognition of the advantages inherent in an interactive view of the relationship between lexical awareness and learning to read. Three somewhat diverse methods of facilitating pre-readers' appreciation of the relationship between spoken and printed language are described. This chapter and the many issues it raises are likely to stimulate considerable research.

A review of research on those children who learn to read prior to attending school stresses the value of observing the ways in which such children teach themselves how to read. Whilst eleven American studies are summarised in a table, it is surprising that no mention is made of Professor Clark's work published here in 1976. In drawing together what is known of early readers and of the reading process, interesting ideas concerning methods of encouraging children to read are developed. Helping the learner form useful hypotheses in terms of the three-level cue system described by Goodman is one. Its implementation poses many problems that have yet to be resolved.

The fifth contribution concerns the child's social environment and its relationship to the acquisition of reading skills. Many of the references are somewhat dated and this seems odd when sociolinguists have been so active in this field. The contention that "... very little is known about (linguistic) awareness" is substantially refuted in the earlier third chapter of the book. In trying to cover a very broad canvas the balance between breadth and depth of treatment in this fifth contribution suffers.

A consideration of the 'Rehabilitation of acquired dyslexia in adolescence', whilst arguably tangential to the theme of the book, looked on the surface to be potentially interesting. The six male and three female subjects were treated for 20 weeks by three different approaches, 'Visuokinesthetic', 'Global stimulation' and 'Phonemic analysis'. The qualitative analysis of the results suggest that the third approach showed some promise. With such (understandably) small numbers and no control group, the optimism is understandable but further studies are needed including evaluation of the eclectic approach to treatment suggested by the author. Additionally, the presentation of the rationale underpinning the treatment programmes is far from convincing. This may have been a function of shortage of space, this chapter being only 16 pages in length. Both descriptive detail and discussion failed to do justice to the topic. One characteristic of dyslexia is a difficulty in coping with sequences of graphic symbols. It is ironic that at some stage in the proof reading a reference to an

article by Becker apparently to be published in the year 9167 has eluded detection and appears in the published text.

The book concludes with a lengthy chapter (77 pages) in which a synthesis between the development of thinking and of reading is presented. The aspiration is laudable, the result less than satisfactory in the reviewer's opinion. An interactionist position concerning thinking and reading is undoubtedly tenable but the theoretical basis presented has many weaknesses. The author draws heavily on Piagetian theory and practice and uses material obtained from children during Piagetian-type interviews. These clinical observations are not really able to stand the weight of theorising erected upon them. Empirical evidence concerning the short and long-term effects of the 'Piagetian and language-orientated' programme would have been welcome. Perhaps as a consequence of this final chapter such studies are now ongoing and will eventually appear.

To what extent have the editors achieved the objectives quoted in the initial paragraph of this review? Some, but not all, of the contributors provide systematic and substantive reviews and syntheses of their particular fields. Some, but not all, of the contributions represent advances in either theory or practice. The seven articles varied greatly in length, ranging from 16 to 72 pages. This was less disconcerting than the variation in quality which ranged from the pedestrian to the excellent. Of course one cannot expect the content of seven different contributions to appeal equally to any reader. However, there are academic criteria concerning both content and format that justify the comment on heterogeneity of quality. Perhaps such a range is almost unavoidable given our present state of knowledge.

On the positive side, certain of the chapters are very useful sources for students interested in specific aspects of reading. Many interesting and controversial ideas that merit formal empirical investigation are suggested. The reviewer leaves this book with an enhanced awareness of the difficulty and the importance of asking profitable research questions. Despite its 'curate's egg' characteristic, the book merits the attention of workers interested in this complex field.

PETER D. PUMFREY.

WICKER, A. W. (1979). *An Introduction to Ecological Psychology*. California: Brooks/Cole, pp. xii + 228, p. \$6.95.

If you study people rather than plants, then 'ecological' may be associated in your mind with the Kansas field station, rather than just fields. If so, you will know what to expect in this book: an historical and critical exposition of the 'ecological psychology' movement pioneered by Roger Barker and Herbert Wright some 30 years ago. Within this brief, the book provides an excellent and readable account, which fills a noticeable gap in the market. Having done his doctoral dissertation at Kansas, Wicker provides a well-informed account, broadly very sympathetic though not lacking in criticism. The book tends to fall into two parts. First, an account of the ideas of behaviour settings, and of synomorphy between people and settings, with discussions of early applications and findings such as 'Midwest and its Children' (in reality Oskaloosa, Kansas, where the field station was sited). Second, several chapters on the theory of under-manning and over-manning of behaviour settings, this being the main theoretical component in ecological psychology to date. Wicker discusses the 'Big School, Small School' work, and similar more recent studies which relate size of school to pupil's experience of responsible positions. The theory has also been applied to churches, and service settings such as the Yosemite National Park.

The final chapter is called 'ecological psychology in the future'. To date, ecological psychology has provided a refreshing perspective on real-life behaviour, but one which has run the risk of being rather separate and insular from other developments. Wicker states that "ecological psychology does not present the 'truth' but rather one version or portion of the truth", and makes a good case for its contribution, with other methods and approaches, to many areas of applied psychology. All in all a useful text if you are interested in these ideas, and the section on schools is of particular relevance to educational psychologists. However, it should be noted that other educational work in this tradition, such as that of Shure in nursery schools, Gump on open plan schools, and Kounin on lesson signals, is not reviewed in the present volume.

P. K. SMITH.

PUBLICATIONS RECEIVED

Books for review have been selected from the following
(' n.p.' means ' no price mentioned ').

- ACADÉMIE DE SCIENCES SOCIALES ET POLITIQUES DE LA RÉPUBLIQUE SOCIALISTE DE ROUMANIE (1979). *Revue Roumaine des Sciences Sociales*. Vol. 23, No. 2, July–December 1979. Bucharest, Roumania.
- ARMSTRONG, S. (Ed.) (1979). *Decade of Change*. Surrey: SRHE, pp. iii+103, p. £4.20.
- ARNOLD, P. J. (1979). *Meaning in Movement, Sport and Physical Education*. London: Heinemann, pp. viii+200, p. £4.95, c. £9.50.
- BAR-TAL, D., and SAXE, L. (Eds.) (1980). *Social Psychology of Education: Theory and Research*. Washington: Hemisphere Publishing Co., pp. iii+384, c. \$17.95.
- BERNSTEIN, A. (1980). *The Flight of the Stork*. New York: Dell Publishing Co., pp. 12+220, n.p.
- BLOOM, B. S. (1979). *Alterable Variables—The New Direction in Educational Research*. Edinburgh: SCRE, pp. 2+14, p. £0.85.
- BOWER, G. H. (Ed.) (1979). *The Psychology of Learning and Motivation, Vol. 13*. New York: Academic Press, pp. v+426, c. \$35.00.
- BROWN, G., and DESFORGES, C. (1979). *Piaget's Theory*. London: Routledge and Kegan Paul, pp. 1+196, c. £7.95, p. £3.95.
- BROWN, P., and FAULDER, C. (1980). *Treat Yourself to Sex—A Guide to Good Loving*. London: Penguin Books, pp. 8+185, p. £0.95.
- BULLOWA, M. (Ed.) (1979). *Before Speech—The Beginning of Interpersonal Communication*. London: Cambridge University Press, pp. vii+400, c. £15.00, p. £4.95.
- CLARK, M. M., and CHEYNE, W. M. (Ed.) (1979). *Studies In Pre-School Education*. London: Hodder and Stoughton for SCRE, pp. vi+292, p. £5.95.
- COHEN, L., and MANION, L. (1980). *Research Methods in Education*. London: Croom Helm, pp. 11+328, c. £11.50, p. £5.50.
- CROSSLEY-HOLLAND, K. (1979). *The Exeter Book Riddles*. Harmondsworth: Penguin Books, pp. 7+139, p. £1.25.
- ELKIND, D. (1979). *The Child and Society. Essays in Applied Child Development*. New York: Oxford University Press, pp. viii+304, n.p.
- FURTH, H. G. (1979). *The World of Grown-Ups: Children's Conceptions of Society*. Limerick, Ireland: Thomond Books, pp. x+221, c. \$16.50.
- GARAN, D. G. (1979). *Against Ourselves—Disorders from Improvements under the Organic Limitedness of Man*. New York: Philosophical Library, pp. v+310, c. \$9.75.
- GRUNEBERG, M. M., and MORRIS, D. E. (1979). *Applied Problems in Memory*. London: Academic Press, pp. v+289, c. £12.40.
- HEYNS, B. (1980). *Summer Learning and the Effects of Schooling*. New York: Academic Press, pp. vii+326, c. \$21.00.
- HOGG, J., and MITTLER, P. J. (Eds.) (1980). *Advances in Mental Handicap Research, Vol. 1*. Chichester: John Wiley and Sons, pp. v+311, c. £15.00.
- HYAMS, B. K. (1979). *Teacher Preparation in Australia: a History of its Development from 1850–1950*. Victoria: ACER, pp. 1+161, p. \$12.00.
- THE INTERNATIONAL UNION FOR CHILD WELFARE (1979). *International Child Welfare Review, No. 43*. Geneva: IUCW, pp. 3+56, n.p.
- JEFFS, A. J. (1979). *Young People and the Youth Service*. London: Routledge and Kegan Paul, pp. vii+150, c. £7.50.
- KAGAN, J. (1980). *The Growth of the Child—Reflections on Human Development*. Andover: Associated Books, pp. 10+283, p. £3.95.
- KITWOOD, T. (1980). *Disclosures to a Stranger—Adolescent Values in an Advanced Industrial Society*. Henley-on-Thames: Routledge and Kegan Paul, pp. 2+291, c. £11.50.
- KLAUSMEIER, H. J., and ASSOCIATES (1979). *Cognitive Learning and Development: Piagetian and Information-Processing Perspectives*. Cambridge: Ballinger Publishing Co., pp. v+370, n.p.
- KWONG, J. (1979). *Chinese Education in Transition: Prelude to the Cultural Revolution*. Quebec: McGill's-Queen's University Press, pp. x+207, c. \$13.95.
- LANGFORD, P. (1979). *Beyond Piaget—Recent Theories of Concept Development and their Significance for Teaching*. Australia: La Trobe University, pp. 1+110, p. £3.50.
- LAVER, M. (1979). *Playing Politics—Seven Games that bring out the Politician in All of Us!* Harmondsworth: Penguin Books, p. £1.25.

- LEEMING, K., SWANN, W., COUPE, J., and MITTLER, P. (1979). *Teaching Language and Communication to the Mentally Handicapped*. (Schools Council Curriculum Bulletin 8.) London: Evans/Methuen Educational, pp. vi+400, c. £9.25.
- MARKOVA, A. K. (1980). *The Teaching and Mastery of Language*. London: Croom Helm, pp. vii+281, c. £12.95.
- MORGAN, D., and McDOWELL, L. (1979). *Patterns of Residence*. Surrey: SRHE, pp. iv+210, p. £7.00.
- O'NEIL, H. F. (Jr.), and SPIELBERGER, C. D. (Eds.) (1979). *Cognitive and Affective Learning Strategies*. New York: Academic Press, pp. v+327, c. \$21.50.
- ORTONY, A. (Ed.) (1979). *Metaphor and Thought*. Cambridge: Cambridge University Press, pp. vi+501, c. £20.00.
- PIPER, K. (1979). *Curriculum Style and Social Learning*. Victoria: ACER, pp. v+169, p. \$10.00.
- PLUCKROSE, H. (1980). *Children in their Primary Schools*. London: Penguin Books, pp. 14+175, p. £1.25.
- PRIDE, D. (1979). *From Communication to Personal Autonomy*. (Language Skills in F.E., and Training, No. 5.) Bolton: Bolton College of Education, pp. 3+66, p. £1.75.
- RUBINSTEIN, D. (1980). *Education and Equality*. London: Penguin Books, pp. 8+335, p. £2.50.
- SHELLARD, J. S. (Ed.) (1979). *Educational Research for Policy Making in Australia*. Victoria: ACER, pp. v+117, n.p.
- SOLNIT, A. J., EISSLER, R. S., FREUD, A., KRIS, M., and NEUBAUER, P. B. (1979). *The Psychoanalytic Study of the Child* (Vol. 34). London: Yale University Press, pp. v+645, c. £15.95.
- STRONG, P. M. (1979). *The Ceremonial Order of the Clinic Parents, Doctors and Medical Bureaucracies*. London: Routledge and Kegan Paul, pp. ix+267, c. £8.95.
- STUART, F. (1979). *Leisure for Mentally Handicapped People*. London: NSMHC Centre, pp. 3+11, p. £0.50.
- TAYLOR, C. (1980). *The Explanation of Behaviour*. Henley-on-Thames: Routledge and Kegan Paul, pp. vii+277, p. £3.75.
- VANDENPLAS-HOLPER, C. (1979). *Education et développement social de l'enfant*. Paris: Presses Universitaires de France, pp. 6+239, n.p.
- VELLUTINO, F. R. (1979). *Dyslexia—Theory and Research*. London: MIT Press, pp. viii+427, c. £10.85.
- WARD, G. (1979). *Learning Time and Teaching For Mastery*. (Occasional Paper No. 15.) Victoria: ACER, pp. v+62, n.p.
- WHITEHURST, G. J., and ZIMMERMAN, B. J. (Eds.) (1979). *The Functions of Language and Cognition*. London: Academic Press, pp. v+313, c. £13.65.

METHUEN



ARTHUR R JENSEN BIAS IN MENTAL TESTING

Anyone claiming that the widely used standardized tests of mental ability — IQ, scholastic aptitude, and achievement tests are culturally biased will have to contend with this book.

Such tests have been criticized by psychologists and journalists, and widely considered unfair to racial minority groups. Twenty-five years ago, Arthur Jensen too believed that standard IQ tests were 'grossly biased against virtually everyone but the white middle-class'. Now, in the most detailed and thoroughly researched book ever published on this topic, Professor Jensen examines psychometric methods for objectively detecting bias in mental testing and for applying standardized tests fairly in education, personnel placement, and other areas.

800 pages Hardback: 0 416 83230 X: £15.00

Third Edition WINFRED F HILL LEARNING A Survey of Psychological Interpretations

Review of the second edition:

'It provides a readable and well-presented introduction to the major theories of learning beginning with Watson's version of behaviourism.' *The Times Higher Education Supplement*

Since this book was first published, learning theory has changed in a variety of ways, some predictable at the time, others unexpected. All these changes are reflected in the third edition. Recent cognitive theories are discussed, and the section on verbal learning and memory has been expanded in a cognitive direction.

362 pages

Hardback: 0 416 73960 1: £7.50

Paperback: 0 416 73970 9: £3.95

The British Journal of Social and Clinical Psychology

JOURNAL TO SPLIT 1981 • JOURNAL TO SPLIT 1981 • JOURNAL TO SPLIT 1981

Pressure of good submissions and an increasing interest in broadening the scope of each part of the journal have persuaded The British Psychological Society to publish two separate journals. In 1981 you will receive these two journals for your single subscription

The British Journal of Social Psychology

The British Journal of Clinical Psychology

You - the subscriber - will then have a year in which to judge the merit of a future subscription to both journals when they become completely independent in 1982. We hope you will take this opportunity to read, re-assess and submit fresh papers to these expanding journals and we will look forward to receiving your ideas and comments

JOURNAL TO SPLIT 1981 • JOURNAL TO SPLIT 1981 • JOURNAL TO SPLIT 1981

Volume 19 (1980) £33.00 (US\$74.00). Single parts £10.00 (US\$22.00)

Volume 18 (1979) £25.00 (US\$57.00)

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN, UK

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

available from the Society's offices

St Andrews House, 48 Princess Road East, Leicester, LE1 7DR

Careers in Psychology	50p
Postgraduate Qualifications and Courses in Psychology	40p
Postgraduate Careers in Clinical Psychology	15p
Postgraduate Careers in Educational Psychology	15p
Postgraduate Careers in Occupational Psychology	15p
Postgraduate Careers in The Prison Service	15p
Postgraduate Careers in Social Psychology	15p

PRESIDENTIAL ADDRESSES

1968 Dissonance and Stress in the Teaching of Psychology: Some Reflections of an Accidental Psychologist: Professor George Seth	10p
1969 Changing Horizons: An Essay in Autobiography: Dr Boris Semeonoff	10p
1970 Choosing: Professor R. J. Audley	10p
1971 In Search of an Idiographic Psychology: Professor H. Gwynne Jones	10p
1972 Psychology Today and Tomorrow: Professor Harry Kay	10p
1973 Psychology in Society: Ends or End: Professor Max Hamilton	10p
1974 On Taking Sides: Professor B. M. Foss	10p
1975 Thought and the Brain: Professor O. L. Zangwill	10p
1976 Psychology and Social Policy: Professor Jack Tizard	10p
1977 The Scientific/Applied Debate in Psychology: A Contribution: May A. Davidson	20p
1978 Predicting Human Development: Professor A. D. B. Clarke	20p

C. S. MYERS LECTURES

1964 Remembering Dr Myers: Professor F. C. Bartlett	10p
1965 Personality and Experimental Psychology: Professor H. J. Eysenck	10p
1967 Practice at liking: A study in experimental aesthetics: Professor D. W. Harding	10p
1968 What is potential ability?: Professor P. E. Vernon	10p
1970 In defence of empirical psychology: Dr D. E. Broadbent	10p
1971 C. S. Myers in retrospect: Professor Alec Rodger	10p
1973 On applied psychology in Britain: Professor George Drew	10p
1975 Measuring psychiatric morbidity in a South Indian population: Dr G. M. Carstairs	10p
1976 Quantitative subjective assessments are almost always biased: Sometimes completely misleading: Dr E. C. Poulton	10p
1977 The neuropsychology of anxiety: Dr J. A. Gray	10p

REPORTS TO GOVERNMENT DEPARTMENTS

To the Royal Commission on the Law Relating to Mental Illness and Mental Deficiency 1954-1957	10p
To the Departmental Committee on the Adoption of Children	10p
To the Butler Committee on the Law relating to the Mentally Abnormal Offender	10p
Comments on the DHSS Review of the Mental Health Act 1959: Professional Affairs Board: <i>Bulletin</i> (1977), 30, 223-226	10p
Evidence to the Data Protection Committee (Home Office): Professional Affairs Board: <i>Bulletin</i> (1977), 30, 224-248	10p
Comments on the Report of the Local Committee on Child Health Services: Professional Affairs Board: <i>Bulletin</i> (1978), 31, 157-160	10p
Evidence to the DHSS Working Group in Primary Health Care Teams: Professional Affairs Board	25p
Comments on DHSS Discussion Document "Marriage Matters": Professional Affairs Board	25p
Comments on Interim Health Building: Note 36, Health Centres: Professional Affairs Board	25p
Comments on the Report of the Committee of Enquiry into Mental Handicap Nursing and Care (Jay Report): Professional Affairs Board	25p
Comments on the Report of the Committee of Enquiry into Normansfield Hospital: Professional Affairs Board	25p
Comments on Home Office Green Paper "Youth Custody and Supervision": Professional Affairs Board	25p

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

(continued)

REPORTS TO GOVERNMENT DEPARTMENTS—continued

Evidence to the Royal Commission on the National Health Service: Professional Affairs Board: March 1977	10p
Comments on the Draft Document on Prevention in the Child Health Services: Professional Affairs Board: May 1978	10p
Evidence to the Home Office Review Committee on the Prison Psychological Service: Professional Affairs Board: May 1978	10p
Comments on the draft health circular on Health Centres and Other Provisions for Primary Health Care: Professional Affairs Board: May 1978	10p
Comments on the DHSS Document on Health Service Complaints Procedure: Professional Affairs Board: August 1975	25p
Comments on "A Happier Old Age": A discussion document from the DHSS and Welsh Office: Professional Affairs Board: September 1978	25p
Manpower Survey: Psychologists and Mental Handicap: Professional Affairs Board: November 1978	25p
Comments on N.D.G. Pamphlet No. 5: Professional Affairs Board: November 1978	25p
Evidence to the Royal Commission on Criminal Procedure: Professional Affairs Board: November 1978	25p
Response to the Consultative Document on the Warnock Report: Professional Affairs Board: March 1979	25p
Evidence to the DHSS Working Party on Observation and Assessment: Professional Affairs Board: 1979	25p

OTHER REPORTS

Principles governing the Employment of Psychological Tests and Clinical Instruments	10p
Psychological Tests: A Statement by The British Psychological Society	10p
Technical Recommendations for Psychological and Educational Tests	10p
The changing role of the Educational Psychologist in Child Guidance Services in Scotland	10p
Teaching Psychology to Nurses (DCP Report)	10p
The Work of the Conference Programme, by Dr George Delafield	10p
The Teaching of Psychology in Schools: Joint Statement by The British Psychological Society and the Association for the Teaching of Psychology: September 1977	15p
British Psychological Education: Scientific Affairs Board: November 1977	20p
Ethical Principles for Research with Human Subjects: Scientific Affairs Board: April 1978	15p
Comments on the Proposed New CCETSW Training Programme: Professional Affairs Board: May 1978	25p
Report of the Working Party on the Sex Discrimination Act: Professional Affairs Board: May 1978	25p
Report of the Working Party on Behaviour Modification: Professional Affairs Board: September 1978	25p
Report of the Working Party on Abuses of Psychology for Political Purposes: Professional Affairs Board: October 1978	25p
Notes for guidance in planning short courses on Psychological Testing: Standing Committee on Test Standards: February 1979	25p
Postgraduate training in Clinical Psychology: Optimal length and format of training: Professional Affairs Board: <i>Bulletin</i> (1977), 30, 284-285	25p
Computer Use in British Psychology Departments: David Hale for Scientific Affairs Board: <i>Bulletin</i> (1977), 30, 410-413	25p
Repertory Grid Methods: Mathematical & Statistical Psychology Section	50p
Monitoring Financial Support for Psychological Research: Scientific Affairs Board: <i>Bulletin</i> (1978), 31, 3-5	25p
Psychological Services for Children: Professional Affairs Board: <i>Bulletin</i> (1978), 31, 125-127	15p
Report of Working Party on Animal Experimentation: Scientific Affairs Board: <i>Bulletin</i> (1979), 32, 44-52	25p
Comments to Royal College of Psychiatrists on the Role of the Community Psychiatric Nurse: Professional Affairs Board: July 1979	25p
Criteria for the Assessment of Postgraduate Training Courses in Clinical Psychology: Professional Affairs Board: November 1979	25p

price includes postage

The British Journal of Medical Psychology

Volume 53, Part 2, June 1980

Contents

- F. N. Watts.** Clinical judgement and clinical training
T. Trauer. Correlates of patient participation in the large group meetings of a therapeutic community
M. L. G. Shaw. The analysis of a repertory grid
D. Childs & R. Hedges. The analysis of interpersonal perceptions as a repertory grid
N. D. Macaskill. The narcissistic core as a focus in the group therapy of the borderline patient
S. P. Llewelyn. The uses of an eclectic approach
P. D. Evans & D. G. White. Sex differences in the development and projection of a simple fear
D. Moss. Transient paranoia and the 'gift' in terminal illness
J. Hayworth, B. C. Little, S. Bonham Carter, P. Raptopoulos, R. G. Priest & M. Sandler. A predictive study of post-partum depression
R. Nelson-Jones & P. Coxhead. Neuroticism, social desirability and attributions affecting self-disclosure
J. J. Ray & R. Bozek. Dissecting the A-B personality type
Book reviews

Volume 53 (1980) £28.00 (US\$62.00). Single parts cost £9.00 (US\$20.00)
Volume 52 (1979) £21.00 (US\$45.00)

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN, UK

The British Journal of Psychology

Volume 71, Part 2, May 1980

Contents

- M. W. Eysenck & H. J. Eysenck.** Mischel and the concept of personality
A. W. Staats. 'Behavioural interaction' and 'interactional psychology' theories of personality
M. O'Mahony & T. Tsang. A preliminary comparison of Cantonese and American-English as taste languages
J. St B. T. Evans. Current issues in the psychology of reasoning
B. B. Hart. Speech modification in near-mute schizophrenics
J. Wilding & N. Mohindra. Effects of subvocal suppression, articulating aloud and noise on sequence recall
M. W. Eysenck & M. C. Eysenck. Effects of processing depth, distinctiveness, and word frequency on retention
G. S. C. Cheong. Relations among symmetry, asymmetry, perceptual comprehension of numerals by kindergarten and first grade children
D. P. Spence & D. R. Beyda. Heart-rate change as a measure of verbal storage and retrieval
S. Folkard & T. H. Monk. Circadian rhythms in human memory
Book reviews

Volume 71 (1980) £35.00 (US\$77.00). Single parts cost £11.00 (US\$24.00)
Volume 70 (1979) £26.00

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts. SG6 1HN, UK

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
Journal of Occupational Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1980 subscription prices, payable in advance, are £35.00 (US \$77.00 in the USA and Canada) for *The British Journal of Psychology*; £33.00 (US \$74.00) for *The British Journal of Social and Clinical Psychology*; £28.00 (US \$62.00) for *The British Journal of Medical Psychology* and £25.00 (US \$55.00) for *Journal of Occupational Psychology*.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £25.00 per volume post free (US \$55.00 in the USA).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK. A special price of £130 (\$290) for all five journals going to the same address is offered for 1980.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £12.00 a year, post free, or £5.00 each number. The prepaid subscription price in the USA is \$25.00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

MANUSCRIPTS

For *The British Journal of Psychology* should be sent to Professor M. Coltheart, Department of Psychology, Birkbeck College, Malet Street, London WC1E 7HX.

For *The British Journal of Medical Psychology* should be sent to Professor J. P. Watson, Department of Psychiatry, Guy's Hospital Medical School, London Bridge, London, SE1.

For *Journal of Occupational Psychology* should be sent to Professor P. Warr, MRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For *The British Journal of Social and Clinical Psychology* should be sent to G. M. Stephenson (Social Psychology Editor), Professor of Social Psychology, Eliot College, University, Canterbury, Kent, CT2 7NS; or to Dr. D. Shapiro (Clinical Psychology Editor), MRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For *The British Journal of Mathematical and Statistical Psychology* should be sent to Dr. C. Leach, Department of Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

For *The British Journal of Educational Psychology* should be sent to Professor Dennis Child, School of Education, St. Thomas' Street, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2.00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £20.00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £4.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,

VOL. 50, PART 2, 1980

CONTENTS

ARTICLES:

LINDSAY, G. A.	The Infant Rating Scale	97
SIEGAL, M. and BOYES, M. C.	The Comparative Effects of Adult and Peer Group Influence on Children's Moral Judgments	105
ELLERMAN, D. A.	Self-regard of Primary School Children: Some Australian Data	114
MERRY, R.	The Keyword Method and Children's Vocabulary Learning in the Classroom	123
RAVIV, A., BAR-TAL, D., RAVIV, A. and BAR-TAL, Y.	Causal Perceptions of Success and Failure by Advantaged, Integrated and Disadvantaged Pupils	137
OWENS, L. and STRATON, R. G.	The Development of a Co-operative, Competitive, and Individualised Learning Preference Scale for Students	147
FRY, P. S. and COE, K. J.	Achievement Performance of Internally and Externally oriented Black and White High School Students under Conditions of Competition and Co-operation Expectancies	162
McKEACHIE, W. J., LIN, Y-G., DAUGHERTY, M., MOFFETT, M. M., NEIGLER, C., NORK, J., WALZ, M. and BALDWIN, R.	Using Student Ratings and Consultation to Improve Instruction	168
LAWSON, M. J. and FUELOEP, S.	Understanding the Purpose of Strategy Training	175

RESEARCH NOTES:

YOUNG, G. M.	The Attitudes of Teachers in Further Education	181
GARDNER, P. L. and TAYLOR, S. M.	A Transmission-Interpretation Scale	186
MILLAR, C. and MACKAY, C. K.	Conflict between Cues in Number Conservation Tasks	188

BOOK REVIEWS	192
------------------------	-----

PUBLICATIONS RECEIVED	207
---------------------------------	-----



THE BRITISH
JOURNAL OF
EDUCATIONAL
PSYCHOLOGY

NOVEMBER 1980

VOL. 50, PART 3

THE BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY

Edited by
Professor DENNIS CHILD

Assistant Editors:

Professor S. N. BENNETT

Dr. G. BROWN

Professor K. WEDELL

Sub-Editor:

Dr. S. CUNNINGHAM

Advised by the Editorial Committee:

Dr. H. BELOFF

Professor R. A. KENNEDY

Professor P. LEVY

Professor J. D. NISBET

Professor E. A. PEEL

Mr. P. PRINCEP

Professor B. SHACKEL

Dr. D. SHARPLES

Professor M. SUTHERLAND

Vol. 50, Part 3

The British Journal of Educational Psychology is issued by the British Psychological Society once a term, in February, June and November. The subscription is £12.00 a year post free, payable in advance, or £5.00 each number.

Subscriptions and Orders should be sent to Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX. Members of the British Psychological Society receive the *Journal* on special terms.

Papers for publication should be sent to Professor Dennis Child, School of Education, St. Thomas' Street, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU. Guidance to contributors on the preparation of papers for the *Journal* is given at the end of the November number each year.

All Books for Review for the *British Journal of Educational Psychology* should be sent to: Dr. G. Brown, Department of Educational Research, The University of Lancaster, Bailrigg, Lancaster.

Advertisements for this *Journal* should be sent to Scottish Academic Press.

Published by

SCOTTISH ACADEMIC PRESS, 33 MONTGOMERY STREET, EDINBURGH, EH7 5JX,
for the British Journal of Educational Psychology, Ltd.

All rights reserved

© 1980. The British Journal of Educational Psychology

PRINTED IN GREAT BRITAIN BY CLARK CONSTABLE LTD., EDINBURGH

PARENTAL INVOLVEMENT AND READING ATTAINMENT

By JENNY HEWISON* AND J. TIZARD

(Thomas Coram Research Unit, Institute of Education, London University)

SUMMARY.—The relationship between a number of home background factors and reading ability was studied in three samples (samples of 63, 100 and 104) of working-class children aged 7-8, using standardised tests given to the children and parental interviews. The home background factor which emerged as most strongly related to reading achievement was whether or not the mother regularly heard the child read ('coached'). WISC IQ scores were obtained on one of the samples ($N = 100$), and it was established that IQ differences did not account for the superior reading performance of the coached children. Controlling for maternal language behaviour, as assessed using scales devised by Bernstein's colleagues, had little effect on the association between coaching and reading performance; controlling for coaching, however, markedly reduced the correlation between maternal language behaviour and reading performance. When the amount of coaching which the children had received was related to reading test score, a highly significant positive association was found. The lack of attention paid in the past to parental involvement in children's school work is commented upon in discussion of the implications of the findings.

INTRODUCTION

THE research reported here is concerned with the influence of parental help on the reading attainments of 7- and 8- year-old children. The work forms part of a more wide-ranging investigation by Hewison (1979) into relationships between the child's home background and his success in school. It is well established that scores obtained by children on tests of academic attainments are strongly associated with demographic characteristics, as measured by social class, material circumstances and size of family—also that attitudinal factors such as parents' newspaper reading habits and the number of books in the home are related to children's educational success (Douglas, 1964; Davie *et al.*, 1972). The aim of Hewison's (1979) study was to investigate whether differences in school achievement *within* a working-class population could be related to differences in the home backgrounds of the children; demographic, cultural and attitudinal variables were to be taken into account, but the main concern was to be with characterising parental *behaviour* and day-to-day child-rearing activities.

Children aged 7-8 years were chosen for study, because that is the age of transfer from infants to junior school and because social class differences in achievement are already well established by that age. Reading attainment was chosen as the measure of school achievement in the studies because at 7-8 years of age it is not only the best index of how well a child is currently doing in school but also the best single predictor of subsequent school achievement.

The pilot inquiry and the two subsequent studies were all carried out in a very homogeneous working-class area, which contained a negligible proportion of families of immigrant origin. All of the families included in the studies were white, and there was little variation in their material and social conditions. In all three studies, information was obtained from reading tests given to the children, and parental interviews.

THE PILOT STUDY

Method

The pilot study was carried out in two infant schools. An attempt was made to interview the mothers of all children in final-year classes at the two schools. Intro-

* Now at the Department of Psychology, University of Durham.

ductory letters were sent to the homes, then an interviewer (J.H.) called at each house and asked the mother if she would agree to be interviewed. Eighty-one homes were contacted in this way. Three mothers did not wish to participate in the study, and seven repeatedly broke interview appointments. Incomplete information was obtained from six homes, where the interviewee was not the full-time guardian of the child. Thus the size of the final interview sample was 65.

In the pilot study structured open interviews were used to obtain the home background information. Topics covered included attitudes to children's play and discipline, the sharing of activities and conversation, reading to the child and hearing him read, a series of questions about how the child spent his leisure time and another on attitudes to school. The time taken to carry out the pilot interview ranged from 30 minutes to two and a half hours.

The reading attainment of the children was assessed using the Southgate Reading Test I. Reading test scores were not obtained on two children who were repeatedly absent from school, so the final sample size for the pilot study was 63. The study took place towards the end of the school summer term, i.e., in June and July when the age range of the children in the sample was 6 : 11 to 7 : 09.

Results

In general, Bernstein's (Bernstein and Brandis, 1970; Brandis, 1974) and others' findings of modest correlations between aspects of the child's behaviour and circumstances at home and his level of attainment were confirmed. There was, however, one question to which replies correlated strongly with the children's reading attainments. This was whether or not the mother regularly heard the child read. Most of the other significant findings between reading ability and parental practices could be accounted for statistically by the fact that parents who displayed attitudes and behaviour which appeared to favour the development of reading ability in their children were also more likely to have the habit of hearing the child read ('coaching'). Analysis by partial correlation showed that, for example, the association between 'attitudinal' factors and reading attainments came about because the 'attitudes' were correlated with 'coaching'—children whose parents had good 'attitudes' but who did not receive 'coaching' were not advanced in their reading.

THE MAIN STUDY

METHOD

The results of the pilot inquiry were both striking and unexpected. In the second study of the series, the ideas generated in the pilot were put to the test on new data.

The names of 30 children were randomly drawn from the first year roll in each of four junior schools. As in the pilot study, introductory letters were sent to the parents of the selected children before an interviewer (J.H.) visited them at home. Once contacted, no mother refused to be interviewed, and none persistently broke appointments: interviewing was terminated when a minimum of 25 successful interviews had been carried out in each school area, 107 interviews being carried out in all.

A closed interview schedule, based on the framework of response coding developed in the pilot study, was used to obtain the home background information. Topics covered in the interviews again included attitudes to school and parental help with reading at home. Aspects of the mother's language behaviour were assessed using two scales devised and employed by members of Bernstein's research team (Brandis, 1970; Henderson, 1971; Brandis, 1974). The first of these scales was designed to assess a mother's willingness to chat to her child in a number of different physical circumstances, and the second her willingness to answer a variety of different 'awkward' questions. The time taken to carry out each interview ranged from 30 minutes to one hour. Scores on a standardised reading test, (the NFER Test 'A')

and on the WISC were obtained for the sample children. The IQ tests were administered by trainee educational psychologists from the London Institute of Education. The testers were not informed about the purpose of the main investigation and since their work was carried out under the supervision of their course tutor, the results are probably as valid, and the testing as reliable as is normally the case.

One child was absent from the reading testing, and six from the IQ testing. The final sample size for whom full reading, IQ and home environment information was available, was therefore 100. The study took place in the autumn term, when the age range of the sample children was 7 : 02 to 8 : 02.

RESULTS

Significant correlations were again obtained between a number of different aspects of the home environment and reading performance. As in the pilot study, the factor which was found to be most strongly associated with reading success was whether or not the mother regularly heard the child read—not whether she read to the child but whether she heard the child read. Taking this factor into account substantially reduced the correlations of the other home environment variables (including whether or not the mother read to the child) with reading performance.

Analysis of variance calculations revealed no significant differences in reading performance among the four schools in the study, nor between boys and girls. No significant school or sex differences were found in the proportions of mothers who heard their children read, and no interactions were found between these variables and the 'coaching' factor in the determination of reading performance, i.e., the 'coaching' effect was observable in all four schools, and for both boys and girls.

Reading attainments and 'coaching'

Of the 100 children in the sample, 47 had mothers who heard them read at home ('coached') and 53 had not. Table 1 presents the reading attainments of the two groups, tested on the NFER Reading Test A.

As in the pilot study, the differences in reading scores between the coached and not-coached samples was very great, amounting to almost one standard deviation. The point biserial correlation between Coaching and Reading Score was $r = 0.61$, i.e., 36 per cent of the variance in reading scores would be accounted for statistically by the factor of coaching.

Reading attainments, coaching and mother's language behaviour

The interrelationship of these three variables is of particular theoretical importance. Mothers' scores on both the language behaviour scales were found to be significantly correlated with their children's reading ability. (Pearson's $r = 0.23$, $P < 0.05$ for the 'Chat' scale; and $r = 0.33$, $P < 0.001$ for the 'Awkward questions' scale.) However, as in the pilot study, it was found that mothers who had educationally favourable language behaviour were also more likely to hear their children read. When this was taken into account, only a very weak and statistically non-

TABLE 1
AGE STANDARDISED READING TEST SCORES OF CHILDREN WHO WERE COACHED OR NOT COACHED

Sample	N	Mean	SD
Coached	47	101.7	10.35
Not coached	53	87.6	7.79
Total	100	94.2	11.46

F ratio = 59.7, df = 1,98, $P < 0.0001$

significant relationship remained between maternal language behaviour and reading performance. (Partial $r = 0.05$ for the 'Chat' scale, and partial $r = 0.16$ for the 'Awkward questions' scale, neither coefficient being significant at the 0.05 level). Controlling for language behaviour, on the other hand, had very little effect on the association between coaching and reading performance. After adjusting for both language variables, partial $r = 0.56$, $P < 0.001$.

Reading attainments, coaching and IQ

As expected, a highly significant relationship existed between reading score and IQ (Pearson's $r = 0.54$, $P < 0.001$). In addition, Coaching and IQ were strongly associated, as shown in Table 2. The point biserial correlation between coaching and IQ was 0.37, $P < 0.001$.

TABLE 2
FULL SCALE WISC IQ SCORES OF CHILDREN WHO WERE COACHED OR NOT COACHED

Sample	N	Mean	SD
Coached	47	108.2	14.59
Not coached	53	97.2	13.04
Total	100	102.4	14.80

F ratio = 16.0, df = 1,98, $P < 0.0001$

Using multiple regression methods, it was established that the IQ differences shown in Table 2 were insufficient to account for the superior reading attainments of the coached children. To demonstrate the inter-relationships of the three variables in another way, groups of children banded for IQ were studied. Table 3 was produced when the banded children were classified by their level of reading attainment (above or below the sample mean of 94.2), and as coached or not coached.

Table 3 shows, first, that the effect of being coached is observable at all levels of IQ; and second, that the higher the IQ band, the higher the proportion of children receiving this form of help with their reading.

TABLE 3
THE RELATIONSHIPS BETWEEN COACHING AND READING PERFORMANCE IN GROUPS OF CHILDREN BANDED FOR IQ

IQ band (Mean IQ = 102.4)		'Poor Readers' Test Score ≤ 94	'Good Readers' Test Score ≥ 95	Total
87 or less	Not coached	10	0	10
	Coached	1	1	2
88-102	Not coached	21	5	26
	Coached	6	10	16
103-117	Not coached	9	3	12
	Coached	2	14	16
118 or above	Not coached	4	1	5
	Coached	2	11	13
Total		55	45	100

A REPLICATION STUDY

Method

A weakness of the studies so far was that no assessment was attempted as to the amount of coaching children were receiving at the time or had received in the past. To

get an idea of this, a further sample of children and their parents was studied. Unfortunately it was not possible this time to assess the children's IQs. The study was a quick one designed simply to find out whether the amount of coaching was related to reading attainment score.

The procedure followed in the main study for obtaining a sample was again adopted. The names of 30 children were randomly drawn from the first-year rolls of the same four junior schools (the study took place 12 months after the previous one) and the contacting and interviewing of parents carried out as before. One hundred and six mothers were contacted, and 105 agreed to be interviewed. One child missed the reading testing, so the final sample size for the study was 104.

Detailed questioning of parents, using a closed interview schedule, made it possible to classify the amount of help their children had received into one of the following five categories: *regular*, or *occasional*, help in the juniors; *regular*, or *occasional*, help in the infants only; *no help* at any time.

Results

Table 4 gives the mean reading scores of children in the five 'coached' categories.

TABLE 4
MEAN READING SCORES OF CHILDREN IN FIVE 'COACHED' CATEGORIES

Help Received	Mean Reading Score	SD	N
Regular help in juniors	104.9	10.9	39
Occasional help in juniors	98.3	8.8	23
Regular help in infants only	91.2	9.3	19
Occasional help in infants only	86.7	9.0	11
No help at any time	81.1	10.3	12
Total sample	96.3	12.8	104

An analysis of variance carried out on the data revealed that the relationship between amount of help given and reading performance was highly significant statistically ($F = 18.5$, $df = 4.99$, $P < 0.0001$).

The results are very consistent indeed. They fully confirm the earlier findings, and suggest in addition that reading attainment scores vary in stepwise fashion with the amount of coaching in reading which the children have received.

DISCUSSION

There is an increasing awareness of the need for co-operation between the home and the school. Also, the significant part that a child's home environment plays in determining his educational progress has been well documented in, for example, the two longitudinal studies carried out by J. W. B. Douglas (Douglas, 1964) and the National Children's Bureau (Davie *et al.*, 1972). Both the Plowden Report (1967) and the Bullock Report (DES, 1975) discussed co-operation between parents and teachers and made constructive suggestions as to how this could be improved.

In general, these recommendations have been concerned with ways of changing the attitudes of teachers and parents. What matters most, according to the Plowden Report is "whether there is a genuine mutual respect, whether parents understand

what the schools are doing for their individual children and teachers realise how dependent they are on parental support ”.

There has been much less discussion of parental involvement in children's actual school work. The Plowden and Bullock Reports both made some reference to parents helping their children with work at home, but neither report made explicit its views as to the desirability or otherwise of this practice. The 'Survey among parents of primary school children', carried out for the Plowden Committee, did, however, yield factual information on the extent to which schools encouraged children to take work home, and on parents' attitudes to homework.

According to their parents' report, 26 per cent of children in the top infants, and 39 per cent in the bottom juniors were given work to do at home. In classes where such work was given, only around 16 per cent of parents disapproved of this, or thought the children were given too much (only 2 per cent thought this); about 60 per cent thought they were given the right amount, and about 20 per cent thought they were given too little (Plowden Report Vol. II, Appendix 3, Table 25). Social class differences in parental response to this question were found to be remarkably small, though significantly more children of parents in social classes I and II were in fact given homework to do (Table 26). Parental help with homework was also found to be widespread, with 73 per cent of children in the top infants receiving help. Less than a third of parents were, however, reported to have asked the school for homework, or how to help their child at home.

The Bullock Report was primarily concerned with the teaching of reading and other language skills in a school setting. Discussion of parental involvement was restricted to general remarks about the role of parents in extending children's language development, and fostering positive attitudes to books and reading. Talking to children and reading to them were considered the most appropriate activities for parents to engage in, and specific warnings were issued against the use by parents of more formal techniques, such as helping their children work through the books in a reading scheme. The Bullock Report did not at any stage address itself to the issues raised by the Plowden inquiry, which had shown that large numbers of children were already being given active help with school work by their parents.

The present study, like the Plowden survey, was concerned with learning more about existing parental practices. It showed, firstly, that many parents (about half, in these samples of white working-class parents) did spend time coaching children in the mechanics of reading, though practically none of these had consulted the school about this, and none at all claimed to have had any encouragement from the school. Secondly, this coaching was undoubtedly related strongly to reading performance. The 'Coaching factor' accounted for 36 per cent of the variance in reading scores—a proportion of variance accounted for which was, incidentally, more than twice the amount of the variance in educational performance within schools accounted for by all variables defined as 'parental attitudes' in the National Survey commissioned by Plowden (Vol. I, Table 1, p. 53.) The effect could be explained, statistically, neither by between-schools differences, nor by IQ effects—the relationship between coaching and reading attainment remaining highly significant even after both these factors had been taken into account. There were no significant sex effects in the data.

IQ and coaching both had independent contributions to make to the prediction of reading performance. This fact established, it remained the case that children who were coached had IQs which were 11 points higher, on average, than those who were not.

The relationship between these two variables may be accounted for in a number of ways which the data cannot distinguish. It may be that more intelligent pupils are more keen to have themselves heard reading by their parents, or that more intelligent

parents have more intelligent children whom they are likely to want to see get on well at school. Another possibility is that coaching which leads to better reading also boosts IQ.

The study leaves unanswered the major question as to why some parents coach and others do not. Those who coached were a self-selected group. It is possible that what was chiefly responsible for their children's success in learning to read was not the practice in reading the parents gave them but rather the interest parents took in their schooling, of which coaching was merely an indicator. The fact that scores on other, 'attitudinal' questions relating to parental interest in the child's education correlated only very modestly with reading attainment score goes against this last hypothesis, but questioning by parental interview may be only a poor indicator of parental 'style' in upbringing.

Survey findings obtained from members of self-defined groups throw no light on the question of how far parental styles of upbringing are subject to change. This can only be decided in practice. Fortunately we had the opportunity, in collaboration with W. N. Schofield, and the local education authority primary advisers to carry out an experiment in the London borough of Haringey. In this study, an effort was made to get *all* parents of children in selected school classes to hear their children read at home. Results of this experiment are in process of analysis and will be presented in subsequent papers.

ACKNOWLEDGMENT.—This paper is based on a doctoral thesis submitted by J. Hewison in partial fulfilment of the requirements for the degree of Ph.D. at the University of London. The empirical work was supported by an MRC postgraduate scholarship, and carried out under the direction of the late Professor J. Tizard. The authors would like to express their gratitude to the Director of Education, London Borough of Barking, for permission to carry out these studies, and to all the head teachers, teachers and parents, for their assistance and co-operation.

REFERENCES

- BERNSTEIN, B., and BRANDIS, W. (1970). Social class differences in communication and control. In BRANDIS, W., and HENDERSON, D. (Eds.), *Social Class, Language and Communication*. London: Routledge and Kegan Paul.
- BRANDIS, W. (1970). Appendix II: A measure of the mother's orientation towards communication and control. In BRANDIS, W., and HENDERSON, D. (Eds.), *Social Class, Language and Communication*. London: Routledge and Kegan Paul.
- BRANDIS, W. (1974). Personal communication.
- DAVIE, R. BUTLER, N., and GOLDSTEIN, H. (1972). *From Birth to Seven: A Report of the National Child Development Study*. London: Longman.
- DEPARTMENT OF EDUCATION AND SCIENCE (1975). *A Language for Life*, (The Bullock Report). London: HMSO.
- DOUGLAS, J. W. B. (1964). *The Home and the School*. London: MacGibbon and Kee.
- HENDERSON, D. (1973). Contextual specificity, discretion and cognitive socialization: with special reference to language. In BERNSTEIN, B. (Ed.), (1973), *Class, Codes and Control: Volume 2*. London: Routledge and Kegan Paul.
- HEWISON, J. (1979). Home environment and reading attainment: a study of children in a working class community. Unpublished Ph.D. thesis, University of London.
- PLOWDEN REPORT (1967). *Children and their Primary Schools*. Report of the Central Advisory Council for Education. London: HMSO.

(Manuscript received 19th July, 1979)

WISC PERFORMANCE SCALE AND COLOURED PROGRESSIVE MATRICES WITH DEAF CHILDREN

By L. EVANS

(Northern Counties School, and honorary lecturer, University of Newcastle upon Tyne)

SUMMARY. The study set out to evaluate the suitability of the WISC Performance test and Raven's Coloured Progressive Matrices for use with deaf children aged 5 to 12 years. The two tests were administered to 125 deaf children, together with a pure-tone measure of hearing loss. Eventual performance in CSE examinations at age 16 was recorded for the older subjects. Test scores were not influenced by degree of hearing loss. The WISC Performance test was found to have good measures of reliability and predictive validity, but presented difficulties in administration to some young or physically handicapped children. The Coloured Progressive Matrices had satisfactory reliability, close comparability with WISC Performance scores and useful predictive validity in the case of the older subjects, but the suitability of this test for use with the younger deaf children was not confirmed.

INTRODUCTION

SEVERE congenital deafness produces communication and language handicaps which have implications for the selection and use of psychological tests. Non-verbal tests are an obvious choice for measuring general ability of deaf children, but as such tests may normally rely upon verbal administration or involve a type of ability possibly related to language competence it should not be too readily assumed that all non-verbal tests are satisfactory or that they yield results with the same meaning as for hearing children. Whilst the use of non-verbal tests with deaf children has proceeded on the basis of their potential suitability, the accumulation of experience and information has indicated a need for further evaluation. This study examined two of the more widely used tests with British deaf children, the Performance Sub-Scale of the Wechsler Intelligence Scale for Children (which will be referred to as WISC Perf) and Raven's Coloured Progressive Matrices (or CPM).

The WISC Perf has been investigated by a number of psychologists working with hearing-impaired children. Murphy (1957) carried out a 're-standardisation' study on a sample of 300 subjects, representing 17 per cent of pupils aged between 6 and 10 years in schools for deaf children in England and Wales at the time. He found a mean IQ of 98 with standard deviation 15.9, which was not significantly different from the norm for hearing children in Wechsler's (1949) original standardisation sample. Subsequent studies by Goetzinger and Rousey (1957) and Brill (1962) in the United States and by Hine (1969) in Great Britain added further findings of deaf or partially hearing children scoring within normal limits.

In a study of reliability of the WISC Perf over a period of time (Evans, 1966) the test was applied to deaf children aged between 6 and 12 years and repeated with the same subjects after an interval of three years. There was close consistency between the mean IQs on the two occasions, 97 and 98 respectively. A high correlation of $r = 0.98$ ($P < 0.01$) between the two sets of scores indicated good test-retest reliability.

Although these findings suggest that the WISC Perf is suitable for use with deaf children, practical experience with the test reveals some problems. Young deaf children, particularly those of low ability, tend to have difficulty in understanding the task involved in some items, and children with additional physical disabilities are handicapped in sub-tests which are scored on speed or involve motor co-ordination. These difficulties highlight the need for other non-verbal tests, for alternative or supplementary use to the WISC Perf.

The Standard Progressive Matrices, being designed as a non-verbal test of general intellectual ability (Raven, 1938), has been assumed to be of particular value for deaf children (Burke, 1958; Buros, 1965). Experimental studies have, however, cast doubt on the capacity of the test to measure the level of general intellectual ability of deaf children. Ewing and Stanton (1943), Denmark (1953) and Gaskill (1957) in Great Britain and Oléron (1950) in France all found deaf or partially hearing subjects to have inferior mean scores in comparison with the norms for hearing children. In a direct comparison of the two tests used with the same subjects (Evans, 1966), 100 deaf children scored substantially below normal in the Standard Progressive Matrices although within normal limits (with mean IQ 99) in the WISC Perf.

The low performance of deaf children may possibly be a consequence of their poor level of language competence. Smith (1964) reported evidence from factorial analysis of a 'verbal loading' of the Progressive Matrices tests, and Lewis (1968) referred to the use of language in solving the more difficult items. A careful examination of Oléron's (1950) results indicate a relationship between test scores and a language experience variable (age of onset of deafness). The possible influence of language competence on test performance may well enhance the value of the Standard Progressive Matrices as a diagnostic and prognostic tool, but the tendency for deaf children to be retarded in the test detracts from its usefulness as a direct alternative to the WISC Perf as a measure of general ability.

Whereas the Standard Progressive Matrices assesses efficiency for thinking by analogous reasoning, the Coloured Progressive Matrices was designed for use with young children at a lower level of intellectual maturity. Raven (1960) considered the Coloured Progressive Matrices to be especially suitable for use with deaf or physically disabled children. In an early British study Gaskill (1957) used this test with a large sample of deaf children and obtained a reasonably normal distribution of results. The test has since become popular for use with deaf children, but Gaskill's promising lead has not been adequately followed up by experimental study to confirm its potential, and there remained a need for further assessment of this version of the Progressive Matrices.

The present study set out to evaluate aspects of the suitability of the Coloured Progressive Matrices for use with deaf children. This was done by direct comparison with the WISC Perf results for the same children. Although the previous evidence on the reliability of the WISC Perf and its comparability with the norms for hearing children supported its use as a criterion of general ability for this purpose, there was a lack of information on the relationship between WISC Perf results and subsequent progress in school work. This study included the further important aim of providing some new information, for both tests, on validity for predicting future educational attainment.

METHOD

Subjects

The subjects were the pupils of a school for deaf children within the age range 5 to 12 years at the time of original testing. There were 125 prelingually deaf children, 70 boys and 55 girls. This could not be regarded as a random sample with regard to distribution of abilities. Apart from a majority of severely deaf children, there were 16 children with moderate hearing loss who were placed at the school for the deaf rather than at partially hearing units on account of slow development, possibly associated with low ability. These subjects were evenly distributed throughout the age range. On the other hand, there were missing from the sample those deaf children of high ability who had transferred from the school for the deaf to partially hearing units after making very good progress. Three such able children had transferred in the year previous to the commencement of the investigation. For these reasons, the distribution of

abilities in the sample could be expected to be biased. The sample also included 12 children with aetiology of brain injury and physical handicaps which could possibly affect their performance in tests involving motor co-ordination. Two-thirds of the additionally handicapped children were in the lower age range, 5 to 9 years. All the children were included in the study, as it was important to gain information on the use of the tests with the less able or additionally handicapped children, but this emphasised the need to compare Coloured Progressive Matrices results with the WISC Perf, as a criterion test, rather than with the norms for hearing children.

Procedure

The sub-tests of the WISC Perf (Block Design, Picture Completion, Picture Arrangement, Object Assembly and Coding) and the book form of the Coloured Progressive Matrices (Sets A, Ab and B) were administered to all subjects during the first year of the investigation. The test instructions were given through an appropriate combination of speech, fingerspelling and signing, to meet the communication needs of individual children. In addition, pure-tone hearing loss was measured. The test results were analysed to provide measures of reliability, comparison of frequency distributions and intercorrelation between the two tests. The test results were also correlated with degree of hearing loss.

For estimating the predictive validity of the tests, eventual performance in the Certificate of Secondary Education was selected as an appropriate objective criterion of educational attainment. Subjects aged 12 years at the time of original testing completed their upper school course and were eligible to sit the CSE examination four years later. As there was only a small number of subjects in each age level, it was necessary to accumulate data over a period of time. This was done for four successive years, at the end of the fifth, sixth, seventh and eighth years of the investigation, to provide data on 52 children at the age of 16 years (who remained at the school, out of the 57 in the middle school age range 9 to 12 years at the original testing).

The actual criterion of successful progress was based upon passes at Grade 4 or better in English, History and Geography. As a Grade 4 pass is, by definition, the average level of attainment for 16-year-old hearing children, and as these subjects have a considerable 'language' involvement, this criterion can be taken as representing an exacting level of attainment for prelingually deaf children. The subjects were placed into two groups according to whether or not they attained this level of success. These were described as a 'CSE' group ($N = 19$) and a 'Non-CSE' group ($N = 33$). This criterion of educational attainment was correlated with the results of the intelligence tests. (As all subjects reached the examination stage at the age of 16 years, it was not necessary to partial age out in the correlations.)

In the statistical analysis it was necessary to use several correlation methods, due to the different types of distributions encountered. The Pearson product-moment coefficient (r) was used for two continuously measurable variables, the biserial (r_b) when one variable was reduced to two categories, the point-biserial (r_{pb}) when one variable was treated as a genuine dichotomy and the tetrachoric (r_{ct}) when both variables were dichotomous. Internal consistency reliability was estimated by correlating raw score totals for odd and even items and applying the Spearman-Brown correction for whole test reliability. As Coding, being a speed test, does not lend itself to the split method, this sub-test score was excluded from the computation of the WISC Perf reliability.

RESULTS

A check was first made to discover whether test scores were influenced by degree of hearing loss. The correlation between hearing loss (average loss for 0.5K, 1K and 2K Hertz in the better ear) and the WISC Perf was $r_b = 0.09$ ($P > 0.05$) and between

hearing loss and Coloured Progressive Matrices was $r_{\cos-p1} = -0.04$ ($P > 0.05$). The correlation with the criterion of educational attainment was $r_{\cos-p1} = 0.23$ ($P > 0.05$). As none of these coefficients was statistically significant, it was not necessary to take further account of hearing loss in the subsequent analysis of the test scores and their influence on educational progress.

In comparing the two tests, the scores were examined in four age groups, except for the comparison of frequency distributions which was restricted to two age groups in order to maintain reasonable numbers of cases in each ability level. As no sex difference was recorded, it was not necessary to break down the groups further. The correlation of test scores with educational attainment was confined to the 52 children (of the original 57 in the 9 to 12 age range) who remained at the school to complete their secondary education.

The main results are presented in Tables 1 to 4.

TABLE 1
INTERNAL CONSISTENCY RELIABILITY OF WISC PERF AND COLOURED PROGRESSIVE MATRICES

Age Range	N	Split-half Correlation r_{tt}	
		WISC Perf	CPM
11-12 yrs	26	0.88*	0.92*
9-10 yrs	31	0.91*	0.81*
7-8 yrs	31	0.95*	0.87*
5-6 yrs	37	0.94*	0.73*

* $P < 0.01$

TABLE 2
PERCENTAGE FREQUENCY DISTRIBUTIONS OF WISC PERF AND COLOURED PROGRESSIVE MATRICES RESULTS

IQ range*	<75 V	75-89 IV	90-109 III	110-125 II	75+ I
Age 9-12 (N = 57)					
WISC Perf	12	23	40	21	4
CPM	14	23	45	16	2
Age 5-8 (N = 68)					
WISC Perf	18	25	38	13	6
CPM	31	26	22	18	3

* WISC Perf IQ or equivalent CPM grade

TABLE 3
CORRELATION BETWEEN WISC PERF AND COLOURED PROGRESSIVE MATRICES

Age Range	N	Correlation Coefficient
		r_{pbl}
11-12 yrs	26	0.83*
9-10 yrs	31	0.65*
7-8 yrs	31	0.56*
5-6 yrs	37	0.56*

* $P < 0.01$

The mean WISC Perf IQ for all subjects was 94.8 with standard deviation 17.74. When examined separately, the results for the 16 severely additionally handicapped children yielded a mean IQ of 79.8. For the remaining 109 children the mean IQ was 98.7.

TABLE 4
WISC PERF AND COLOURED PROGRESSIVE MATRICES CORRELATED WITH CSE PERFORMANCE

Age Range†	N	Correlation with CSE†	
		WISC Perf r_b	CPM r_{008-11}
9-12 yrs	52	0.81 [*]	0.70 [*]

* $P < 0.01$

† CSE performance between 4 and 7 years later at age 16 years

‡ age at time of original testing

DISCUSSION AND CONCLUSIONS

The finding of no significant correlations between test scores and the results of the hearing test was in agreement with previous studies which reported lack of relationship between hearing loss and general ability as measured in performance tests of intelligence within groups of deaf subjects (Hood, 1949; Glowatsky, 1953; Kendall, 1953; Murphy, 1957; Pickles, 1966; De Marco, 1970).

The measures of internal reliability were satisfactorily high for the WISC Perf, comparing favourably with the estimates in Wechsler's (1949) original standardisation with hearing children.

The distribution of WISC Perf IQs for the older subjects, in the 9 to 12 years age range, was reasonably close to normal, but with a cluster of very low scores. The distribution for the younger children was more markedly skewed, but this can be explained by the majority of the additionally handicapped children of low ability being in this group. Taking into account the factors already mentioned which influence the level of ability of this school population, these results are not inconsistent with previous findings which support the use of the WISC Perf as a measure of general ability in deaf children.

The degree of correlation between WISC Perf results at middle school age and subsequent success at age 16 in CSE examinations in the 'language' subjects can be regarded as a good level of validity for predicting future attainment over a medium term. There is lack of previous information to make a direct comparison of forecasting efficiency, although Murphy (1957) found WISC Perf results to correlate $r = 0.64$ with teachers' ratings of current academic attainment and Pickles (1966) reported a similar correlation of $r = 0.61$ for slow-learning deaf children.

For the Coloured Progressive Matrices there was a consistent age difference. Internal reliability measures were generally satisfactory, but for the youngest age group the estimate was rather low for a reliability coefficient.

For the 9- to 12-year-old subjects the distribution of Coloured Progressive Matrices scores was closely comparable to the corresponding distribution of WISC Perf IQs, and as such this version was closer to the normal for hearing children than the results of the previous investigations using the Standard Progressive Matrices with deaf children. However, the younger subjects had an abnormal distribution with a

relatively greater proportion of low scores than for the WISC Perf. This suggests the possibility that a minimal amount of linguistic competence is required for the solution of the problems, which is still undeveloped in the younger deaf children.

The Coloured Progressive Matrices scores for the older subjects correlated closely with WISC Perf results. The correlation of $r = 0.83$ for the 11- and 12-year-old subjects exactly confirmed Gaskill's (1957) finding. The degree of correspondence between the scores in the two tests was lower for the younger subjects.

The correlation between the Coloured Progressive Matrices and later CSE performance was, of course, obtained only for the older subjects, aged 9 to 12 years at the time of the original testing. The resultant correlation, although lower than that for the WISC Perf, can be regarded as a useful level of predictive validity.

These results lead to the overall conclusion that the WISC Perf is a suitable test of general ability for use with deaf children, giving reliable results with useful validity for forecasting educational attainment. The test was not completely satisfactory for the 5- and 6-year-old children, and physically handicapped children had especial difficulty with sub-tests involving motor co-ordination or scored on speed. This confirmed the need for a complementary test which, whilst avoiding these difficulties, yields comparable results. The results of the Coloured Matrices suggest that this is a suitable test for this purpose in the case of the children in the age range 9 to 12 years, but do not confirm its suitability for use with younger deaf children.

ACKNOWLEDGMENTS.—The author records his thanks to the staff and pupils of Northern Counties School, for co-operation with the study.

REFERENCES

- BRILL, R. G. (1962). The relationship of Wechsler IQs to academic achievement among deaf students. *Except. Child.*, 28, 315-321.
- BURKE, H. R. (1958). Raven's Progressive Matrices—a review and critical evaluation. *J. Genet. Psychol.*, 93, 199-228.
- BUROS, O. K. (1965). *The Sixth Mental Measurement Yearbook*. New Jersey: Gryphon Press.
- DE MARCO, W. (1970). The scatter of intellectual abilities of the hard of hearing as assessed by the Wechsler Intelligence Scale for Children. *Dissertation Abstr. Int.*, 30, 3383-3384.
- DENMARK, F. G. W. (1953). An investigation of some problems in the education of partially hearing children. Unpublished M.A. Thesis, University of Liverpool.
- EVANS, L. (1966). A comparative study of the Wechsler Intelligence Scale for Children (Performance) and Raven's Progressive Matrices with deaf children. *Teach. Deaf.*, 64, 76-82.
- EWING, A. W. G., and STANTON, D. A. G. (1943). A study of children with defective hearing. *Teach. Deaf.*, 41, 56-59.
- GASKILL, P. (1957). Tests of abilities and attainments. In EWING, A. W. G. (Ed.), *Educational Guidance and the Deaf Child*. Manchester University Press.
- GLOWATSKY, E. (1953). The verbal element in the intelligence scores of congenitally deaf and hard of hearing children. *Am. Ann. Deaf.*, 98, 328-335.
- GOETZINGER, C. P., and ROUSEY, C. L. (1957). A study of the Wechsler Performance Scale (Form II) and the Knox Cube Test with deaf adolescents. *Am. Ann. Deaf.*, 102, 388-398.
- HINE, W. D. (1969). The abilities of children with partial hearing. Unpublished M.A. Thesis, University of Liverpool.
- HOOD, H. B. (1949). A preliminary survey of some mental abilities of deaf children. *Br. J. educ. Psychol.*, 19, 210-219.
- KENDALL, D. C. (1953). The mental development of young deaf children. Unpublished Ph.D. Thesis, University of Manchester.
- LEWIS, M. M. (1968). *Language and Personality in Deaf Children*. Slough: National Foundation for Educational Research.
- MURPHY, L. J. (1957). Tests of abilities and attainments. In EWING, A. W. G. (Ed.), *Educational Guidance and the Deaf Child*. Manchester University Press.
- OLÉRON, P. (1950). A study of the intelligence of the deaf. *Am. Ann. Deaf.*, 95, 179-195.
- PICKLES, D. G. (1966). The Wechsler Performance Scale and its relationship to speech and educational response in deaf slow-learning children. *Teach. Deaf.*, 64, 382-392.

- RAVEN, J. C. (1938). *Progressive Matrices*. London: H. K. Lewis.
RAVEN, J. C. (1960). *Guide to Using the Coloured Progressive Matrices*. London: H. K. Lewis.
SMITH, I. M. (1964). *Spatial Ability*. London University Press.
WECHSLER, D. (1949). *Wechsler Intelligence Scale for Children*. New York: The Psychological Corporation.

(Manuscript received 10th January, 1980)

PRINCIPAL COMPONENTS ANALYSIS OF THE REVISED BRISTOL SOCIAL ADJUSTMENT GUIDES

By P. A. McDERMOTT

(Graduate School of Education, University of Pennsylvania)

SUMMARY. In clinical and research application, the various syndromes and syndrome combinations of the Bristol Social Adjustment Guides (BSAG) are used in the initial identification and subsequent classification of behavioural and social maladjustment in school children. It is important to test the validity of the assumptions that the overall BSAG scales, i.e., under- and over-reaction, retain sufficient factorial integrity for use as general indicators of maladjustment and that the several contributing syndromes evidence a satisfactory degree of specificity to be considered as individual dimensions of maladaptive behaviour. Data for the standardisation sample of 2,527 randomly chosen children for the revised BSAG were submitted to principal components factor analysis with varimax rotation of significant factors. A two-factor solution was found with the resulting factor structure clearly substantiating the validity of an under- vs. over-reactive dichotomy of maladjusted behaviours. Moreover, significant syndromic specificity was established for the majority of syndromes, with conjoint and reciprocal variation being suggested in the case of two of the syndromes. The findings, in general, support the uses of the BSAG as advised by the developer in the test manual and associated research.

INTRODUCTION

THE BRISTOL SOCIAL ADJUSTMENT GUIDES (BSAG), both in the original edition (Stott and Sykes, 1956) and in the revision (Stott, 1974) have gone further than other measures of behaviour disturbance in claiming to identify specific syndromes beyond the well-known dichotomy of inhibited and acting-out behaviours. The main purpose of this article is to report the re-treatment of the data for the revised BSAG in an effort to test the substance of this claim.

The revised BSAG is composed of a collection of 110 brief verbal expressions to which classroom teachers may respond in describing children's typical behavioural styles in the school context. These expressions, having been drawn originally from the terminology employed by teachers in describing children's adjustment difficulties, are largely phenomenologically focused and require little, if any, teacher interpretation or explication of the meaning of observed behaviours. Additionally, the BSAG provides a complementary set of "well adjusted" expressions which teachers may alternatively select in describing a child.

In order to produce a typology of maladjustment problems, the 110 behavioural indicators are gathered into several mutually exclusive and homogeneous groups of behaviours. This grouping process was performed on the basis of a series of non-parametric-like clustering techniques which applied a 'specificity ratio' (Stott *et al.*, 1975) by assigning each indicator to that group of other items with which it was most frequently associated by simultaneous occurrence in the normative population. This procedure resulted in five *core syndromes* of social maladjustment and three non-syndromic *associated groupings* of malbehaviours.

As described by Stott (1974), the core syndromes of similar behaviours are: (a) *unforthcomingness*, a general unwillingness or inability to express and extend oneself assertively in social and learning settings, (b) *withdrawal*, an indifference to or aversion to association with other people, (c) *depression*, a motivational deficit manifesting itself in a failure to respond to, or to seek out, environmental stimulation, (d) *inconsequence*, i.e., impulse-ridden behaviour carried out in the absence of pre-planning or consideration of possible adverse consequences, and (e) *hostility*, a complex of attacking

and avoiding behaviours which operate to sever unrewarding affiliative relationships through provocative and unfriendly postures and behaviours.

The non-syndromic groupings, on the other hand, contain behavioural indicators which are neither homogeneous to one another nor exclusively related to any single core syndrome. They include (a) *non-syndromic under-reaction*, referring to a collection of phenotypically similar non-behaving or passive-like indicators which are not interpretable in themselves but which are corroborative of maladjustment as suggested by unforthcomingness, withdrawal or depression, and (b) *non-syndromic over-reaction*, connoting a set of behaviours relating to delinquent or socially deviant reactions that are not syndromically specific, but rather provide supporting evidence for inconsequential or hostile behaviour. A final associated grouping, called *peer-maladaptiveness*, although believed by Stott *et al.* (1975) to possess a moderate degree of internal item homogeneity, is described as a group of indicators pointing to poor peer relations which may coexist with or evolve from otherwise syndromically specific inconsequential or hostile behaviour.

The various syndromes and associated groupings are further divided into those which are under-reactive and those which are over-reactive in nature. Correspondingly, the raw scores for the unforthcomingness, withdrawal and depression syndromes, as well as for the under-reactive associated grouping, are combined to yield a total *under-reaction* scale score. A similar scale score for general *over-reaction* is produced by adding the scores for inconsequence, hostility, non-syndromic over-reaction and peer-maladaptiveness. In general clinical practice, these two composite scales are frequently considered the principal indices for initially determining whether or not a child is evidencing significant levels of maladjustment. If the score for either the under- or over-reaction scale is significantly deviant from population means, the child may be viewed as socially maladjusted. Thereafter, the particular nature and dynamics of a child's maladaptive style may be investigated and further explained through a search for problems in specific syndromic and associated reactions.

The appropriateness of the behavioural observation content and the discriminatory efficiency of the BSAG in the identification of childhood maladjustment have been fairly well established in a variety of criterion-related and predictive validity research evidence (e.g., Davis *et al.*, 1972; Stott *et al.*, 1975; Stott and Wilson, 1977; Stott, 1978, 1979). Moreover, the construct validity of the multisyndromic approach to defining behavioural maladaptation would appear clinically and intuitively meaningful. However, the factorial integrity and relationships among the several syndromes and associated groupings for the normative population upon which the revised BSAG was based have never been substantiated through factor analytic study. The present research was directed to test the assumption that the behavioural observation data for the latest revision of the BSAG were indeed reflecting a complex of functionally identifiable and specific dimensions of social maladjustment that might therefore lend themselves to a viable typology for clinical differentiation among maladjusted children.

METHOD

Subjects

The standardisation sample for the revised BSAG (Stott, 1974) consisted of 2,527 public and private school children drawn at random from the school population of an urban centre and surrounding rural area of Ontario province. The children ranged in age from five years to approximately 15 years with a total of 1,305 boys and 1,222 girls and a near-equal distribution of age-mates.

Procedure

Each child was independently observed by a classroom teacher who then responded to the various adjustment and maladjustment indicators on the BSAG. There-

after, scores on the 110 maladjustment items for each child were combined to produce a raw score for each of five core syndromes and three associated groupings.

Statistical analysis

Children's scores for each of the resultant syndromes and groupings were intercorrelated. The correlation matrix was then submitted to a principal components factor analysis with resulting factors having an eigenvalue greater than one being rotated according to Kaiser's (1958) varimax criterion.

In order to determine the syndromic specificity of the individual groupings, Cohen's (1959) 'subtest specificity' method was applied. Essentially, this technique proceeds to partial out the several sources of variance found in each syndrome such that variance related solely to a particular syndrome or grouping may be isolated. This is accomplished by subtracting the communality (i.e., the proportion of a syndrome's variance shared in common with other syndromes) for each syndrome from its respective reliability as represented by the alpha coefficient for internal consistency (found in the revised BSAG manual, Stott and Marston, 1971). The quantity obtained is an estimate of a syndrome's true variance, i.e., that proportion of the variance which is both reliable (not due to error of measurement) and distinctive (not shared with other syndromes).

To assess the significance of the specificity for each syndrome, specificity values were compared with estimated error variance, where error variance equals one minus the alpha coefficient. Levels of specificity in excess of estimated error variance were considered significant. This is because it is that specificity component of a syndrome's variance, and only that component, which can carry the variance required to support the idea that a certain syndrome is measuring qualities attributed exclusively to that particular syndrome.

RESULTS

Table 1 presents the correlations among the syndromes and associated groupings. Fairly consistent patterns of moderate positive intercorrelations may be noted among those groupings commonly referred to as under-reactive in style as well as among

TABLE 1

INTERCORRELATIONS AMONG SCORES ON THE CORE SYNDROMES AND ASSOCIATED GROUPINGS FOR THE BSAG STANDARDISATION SAMPLE

Syndrome or Grouping	U	W	D	UR	Q	H	PM
Unforthcomingness (U)							
Withdrawal (W)	22						
Depression (D)	12	46					
Non-Syndromic Under-reaction (UR)	58	41	36				
Inconsequence (Q)	-20	02	19	-03			
Hostility (H)	-10	08	18	02	43		
Peer-Maladaptiveness (PM)	-11	02	14	-01	54	58	
Non-Syndromic Over-reaction (OV)	-14	13	31	02	65	50	50

Note. Decimal points omitted. N = 2,527.

those thought of as over-reactive in style, respectively. The principal factor solution extracted two significant factors cumulatively accounting for approximately 60 per cent of the total variance of scores among the groupings. The matrix of principal varimax rotated factor loadings, communalities and specificity indices is displayed in Table 2.

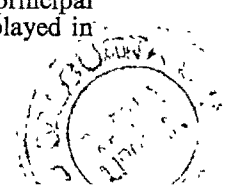


TABLE 2
VARIMAX ROTATED PRINCIPAL FACTOR MATRIX FOR THE REVISED BSAG

Syndrome or Grouping	Factor		Communality	Specificity*
	I	II		
Unforthcomingness	-0.22	<i>0.56</i>	0.36	<i>0.38</i>
Withdrawal	0.10	<i>0.55</i>	0.31	0.28
Depression	0.29	<i>0.49</i>	0.32	0.34
Non-Syndromic Under-reaction	-0.04	<i>0.85</i>	0.73	0.16
Inconsequence	<i>0.75</i>	-0.03	0.56	<i>0.27</i>
Hostility	<i>0.66</i>	0.04	0.44	<i>0.32</i>
Peer-Maladaptiveness	<i>0.70</i>	-0.01	0.50	<i>0.26</i>
Non-Syndromic Over-reaction	<i>0.79</i>	0.08	0.64	0.03
% common factor variance	34.5	26.0		
% cumulative factor variance	34.5	60.6		

Note. Rotated factor loadings ≥ 0.40 are italicised. $N = 2,527$.

* Values for specificity were obtained by subtracting the communality for each syndrome or grouping from its index of internal consistency (as reported in the BSAG manual, Stott, 1974, p. 11). Whereas communality values indicate the proportion of variance a syndrome or grouping shares in common with other syndromes or groupings, specificity values reflect the proportion of variance in the factor structure that is both reliable and unique to a particular syndrome or grouping. Specificity indices are italicised where the value exceeds estimated error variance, i.e., 1 minus the internal consistency coefficient, for each syndrome or grouping.

According to the patterns of significant factor loadings, i.e., those greater than 0.40, the first factor was clearly composed of a combination of the two over-reactive type core syndromes and two associated groupings and was therefore interpreted as a factor of *general over-reaction*, whereas the second factor was founded upon a combination of the three under-reactive type syndromes and one associated grouping, hence constituting a *general under-reaction* factor.

Tests of syndrome specificity demonstrated significant specificity for three of the core syndromes, i.e., unforthcomingness, inconsequence and hostility, and for one of the associated groupings, i.e., peer-maladaptiveness. Specificity was not supported for two core syndromes, i.e. withdrawal and depression. As would have been anticipated in the case of the two non-syndromic groupings (under- and over-reaction which were clearly and deliberately comprised of heterogeneous combinations of behavioural indicators), no specificity was indicated.

DISCUSSION

The identification of two principal orthogonal factors (i.e., general under- and over-reaction) amidst the BSAG's behavioural indicators is confirmatory of the appropriateness of the proposition by Stott *et al.* (1975, pp. 72-75) that the total scores for the respective under- and over-reaction scales be used as primary indices of child maladjustment in nomothetic comparisons. This finding further coincides with results from a variety of other factor analytic studies (Lessing and Zagorin, 1971; Quay, 1972; Spivack and Swift, 1973; Felner *et al.*, 1975) which have supported the concept of major behavioural or personality dimensions of the inactive/withdrawn *vs.* active/provocative order among children. Also Hale (1978), in applying factor analytic procedures to examine the factor structure of the BSAG among a group of rural school children from the midwestern United States, more than half of whom had been referred for psychological evaluation in view of suspected mental deficiency, learning difficulties or other problems, derived a two-factor solution identical to the present one based on the normative sample.

Evidence for the syndromic specificity of unforthcomingness, in consequence and hostility, as well as for the single associated grouping believed to approach some syndromic specificity (Stott *et al.*, 1975, p. 65), i.e., peer-maladaptiveness, is supportive of the idea of basing differentiation and classification upon such groupings. Comparisons of syndromic specificity to error variance were not supportive of the specificity hypothesis in the cases of withdrawal or depression. This is reminiscent of the observations by Stott *et al.*, (1975 pp. 45-51) that, in the original clustering process for grouping the behavioural indicators, indicators assigned to the withdrawal and depression syndromes, respectively, were competing runners-up for reciprocal membership in the opposite syndrome. However, this is not to diminish the typological importance of these latter syndromes. Indeed, there is considerable empirical evidence to the contrary.

In a recent reanalysis of the standardisation data for the BSAG (McDermott, 1980), the 2,527 children were first grouped by a hierarchical cluster analysis process into a number of maximally homogeneous profile groups in terms of similar profiles on the various syndromes. This procedure was conducted to test the hypothesis that, when significant maladjustment occurs among children, it does so usually by manifestation of significant elevations on more than one syndrome. It was found that multisyndromic involvement constitutes 60 per cent of all maladjustment. More specifically, more than 39 per cent of all maladjustment involved some combination of either the withdrawal or depression syndrome with other syndromes or with one another. In a subsequent multiple discriminant analysis to extract the principal linear functions that contributed to the separation of the homogeneous groups of syndromic profile types, it was discovered that withdrawal and depression operate *conjointly* and *reciprocally* to account for nearly 25 per cent of group differentiation among all adjusted and maladjusted children.

In examining the listing of specificity values presented in Table 2, it will be noted that the entries for all five core syndromes and peer-maladaptiveness exceed 0.25 (the average specificity index for these six groupings is 0.31). The 0.25 level was established by Kaufman (1975) as a suitable criterion for moderate specificity of a test. This observation, in company with the discovery that withdrawal and depression covary to distinguish among over one quarter of social adjustment patterns (McDermott, 1980), leads to the surmise that, whereas these syndromes hold no apparent specificity in the population of children as a whole, they may well exhibit significant specificity for some subpopulation of children such as maladjusted children.

To evaluate this possibility, the present author re-examined the principal varimax rotated factor matrix published for the Hale study (1978, Table 2) of the BSAG for a largely clinical sample. The communalities and specificity values were then calculated and compared with estimated error variances. As suspected, the specificity values for both the withdrawal and depression syndromes were markedly significant. The thrust of these findings is to suggest that, depending upon the subpopulation of children being considered, the withdrawal and depression syndromes will either covary positively (conjoint variation as in a mostly normal sample) or negatively (reciprocal variation as in a clinical sample) in their contribution to a typology of social maladjustment.

CONCLUSION

The results are supportive of the concept that children's maladaptive behaviour may be meaningfully associated with either an under-reactive or over-reactive dimension of maladjustment. This would affirm the validity of using overall BSAG under- and over-reaction scores as initial indices of maladjustment. Moreover, the efficacy of viewing the individual groupings of maladjustment indicators as sufficiently

specific for founding a typology upon their variations was confirmed in the case of unforthcomingness, all of the over-reactive core syndromes and peer-maladaptiveness. For the case of the withdrawal and depression syndromes, whereas their individual specificity was not demonstrated for the normative sample as a whole, the present findings considered in the light of other emergent research would raise the possibility of the conjoint and reciprocal variation of these syndromes among children.

ACKNOWLEDGMENTS.—The author wishes to thank Dr. Denis H. Stott, former professor of psychology at the University of Guelph, Ontario, for kindly making the revised Bristol Social Adjustment Guides data available for this study.

Requests for reprints of this article should be sent to Paul A. McDermott, Graduate School of Education, University of Pennsylvania, 3700 Walnut Street, Philadelphia, Pennsylvania 19104.

REFERENCES

- COHEN, J. (1959). The factorial structure of the WISC at ages 7-6, 10-6 and 13-6. *J. consult. Psychol.*, 23, 285-299.
- DAVIS, R., BUTLER, N. R., and GOLDSTEIN, N. (1972). *From Birth to Seven: Second Report of the National Child Development Study*. London: Longman and National Children's Bureau.
- FELNER, R. D., STOLBERG, A., and COWEN, E. N. (1975). Crisis events and school mental health referral patterns of young children. *J. consult. clin. Psychol.*, 43, 305-310.
- HALE, R. L. (1978). A factor analytic study of the Bristol Social Adjustment Guides in a rural school population. *Psychol. Rep.*, 42, 215-218.
- KAISER, H. F. (1958). The Varimax criterion for analytic rotation in factor analysis. *Psychomet.*, 23, 187-200.
- KAUFMAN, A. S. (1975). Factor analysis of the WISC-R at 11 age levels between 6½ and 16½ years. *J. consult. clin. Psychol.*, 43, 135-147.
- LESSING, E., and ZAGORIN, S. (1971). Dimensions of psychopathology in middle childhood as evaluated by three symptom checklists. *Educ. psychol. Meas.*, 31, 175-198.
- MCDERMOTT, P. A. (1980). Prevalence and constituency of behavioral disturbance taxonomies in the regular school population. *J. abnorm. child Psychol.*, 8, in press.
- QUAY, H. C. (1972). Patterns of aggression, withdrawal, and immaturity. In QUAY, H. C., and WERRY, J. S. (Eds.), *Psychopathological Disorders of Childhood*. New York: Wiley.
- SPIVACK, G., and SWIFT, M. S. (1973). The classroom behavior of children: a critical review of teacher-administered rating scales. *J. spec. Educ.*, 7, 55-89.
- STOTT, D. H. (1971). Classification of behaviour disturbance among school-age students: principles, epidemiology and syndromes. *Psychol. in Schools*, 8, 232-239.
- STOTT, D. H. (1974). *Manual to the Bristol Social Adjustment Guides*. London: Hodder and Stoughton.
- STOTT, D. H. (1978). Epidemiological indicators of the origin of behaviour disturbance as measured by the Bristol Social Adjustment Guides. *Genet. Psychol. Monogr.*, 97, 127-159.
- STOTT, D. H. (1979). The British Social Adjustment Guides. *Ther. Educ.*, 7, 34-44.
- STOTT, D. H., and MARSTON, N. C. (1971). *The Bristol Social Adjustment Guides*. San Diego: Educational and Industrial Testing Service.
- STOTT, D. H., MARSTON, N. C., and NEILL, S. J. (1975). *Taxonomy of Behaviour Disturbance*. London: Hodder and Stoughton.
- STOTT, D. H., and SYKES, E. G. (1956). *The Bristol Social Adjustment Guides*. London: University of London Press.
- STOTT, D. H., and WILSON, D. M. (1977). The adult criminal as juvenile. *Br. J. Criminol.*, 17, 47-57.

(Manuscript received 14th March, 1979)

MODE OF STIMULUS PRESENTATION AND SHORT-TERM MEMORY EFFICIENCY IN PRIMARY SCHOOL CHILDREN

By D. FONTANA AND HILARY EVANS
(Department of Education, University College, Cardiff)

SUMMARY. A sample of primary school children was given a range of short-term memory tests to determine whether recall was optimised by aural or by visual stimuli presentation. Results showed that for both sexes and for each age group aural presentation was consistently superior to visual. These results have important potential implications both for theories on the working of short-term memory and for teaching techniques in the primary school.

INTRODUCTION

THE importance of short-term memory for success both in reading skills (e.g., Guthrie and Goldberg, 1972; Torgeson, 1975) and in mathematics (e.g., Basu, 1972; Lunzer *et al.*, 1976) has been stressed by a variety of investigators. However, a number of important questions remain to be solved on short-term memory processes in children. There is as yet, for example, little consistent developmental evidence on the increase in short-term memory span with age, or on the relationship, if any, between short-term memory efficiency and type of stimulus material. In particular, we lack unequivocal findings on the effect that *mode of stimulus presentation* has upon optimum short-term memory performance. Such evidence as is available to us, in fact, is marked by an unusual degree of contradiction, in particular where the relative benefits of visual and auditory presentation of both words and digits are under scrutiny.

Over the last two decades, a number of research findings relevant to the problem have been published, particularly in the USA. Murray and Roberts (1968) found digits presented aurally produced better short-term memory performance in girls aged 7 through 10 than digits presented visually, and that this discrepancy increased with age. Dornbush and Basow (1970) produced similar findings for first, third, fifth and ninth grade pupils, as did Baumeister (1974) with second and sixth grade retarded pupils. However, Katz and Deutsch (1967) though obtaining comparable results for digits, found that visual presentation appeared superior to auditory when the stimulus material consisted of words. By further contrast, Koppitz in a series of investigations (most recently 1977) demonstrated that for her samples (first grade to fifth grade children) *visual presentation was consistently superior to auditory even for digits*, and that again the discrepancy increased with age. Siegal and Allik (1973), using kindergarten, second grade, fifth grade and college students, produced better results for animals and objects presented visually than aurally, while Perlmutter and Myers (1975) and Spitzer (1976) obtained similar findings for kindergarten and for 5-, 9- and 11-year old children.

In suggesting a reason for this contradiction one of the obvious problems is that we have no universally accepted model of the working of short-term memory to which to make reference. Sperling (1963) suggests that all stimuli must be coded acoustically in the short-term store, and if this is the case auditory stimuli are potentially superior to visual in that they require no modality transfer before storage. However, Perlmutter and Myers (1975) consider visual stimuli can be processed in *both* modalities, and since such processing may be simultaneous this eliminates the potential superiority of auditory material, while Koppitz (1977) has it that visual presentation is actually superior because it offers better opportunities for cumulative rehearsal (and thus for retention maintenance) than does auditory. It seems possible in the present

instance that experimental design itself may on occasion have added to the confusion. For example Siegal and Allik (1973) used pictures of animals and objects in their visual presentation strategy but word names in their auditory, while Perlmutter and Myers (1975) used actual objects in the visual mode but again word names in the auditory. This raises doubts on the degree of equivalence between the modalities and thus on the equivalence of the coding tasks facing short-term memory.

A clear need therefore exists to resolve the question of modality superiority, and to resolve it through an experimental design that strives for equivalence of stimuli presentation. The implications of such a resolution for education are clear, particularly in the primary school sphere where the first two statements in the often quoted Chinese proverb, 'I hear and I forget, I see and I remember, I do and I understand', indicate an apparent preference on the part of educators for visual rather than auditory stimuli.

METHOD

Sample

The sample consisted of four year groups (classes) spanning the 7- to 11-year-old range in a local primary school. After eliminating absentees and those nominated by the school as backward readers the full sample consisted of 99 children (52 boys and 47 girls). Breakdown of the sample by year groups and sex, together with details on mean ages, appears at the appropriate points in the tables below.

The tests

The tests used in the experiment consisted of digits and words on the rationale that these are the units of information most frequently encountered by children in their school work. Four tests were constructed, two auditory and two visual, as follows (see Appendix for the tests):

Test 1. Auditory Digit Span. This test consisted of random digits arranged in lists running from two digits (shortest list) to nine digits (longest list), and with no digit appearing more than once on any list. Each list was tape-recorded with one-second intervals between digits, and terminated with a bleep which was the signal for recall. Subjects were presented with lists in ascending length order, and after a trial run were asked to recall each list correctly before proceeding to the next. Where failure occurred they were given a second chance from a reserve list of equal length and tape-recorded in an identical fashion. The test was terminated for each subject when (s)he had twice failed at a given length, and their auditory digit span score was deemed to be the length of the last list satisfactorily recalled.

Test 2. Auditory Word Span. Test 2 followed the same pattern as Test 1 except that the material consisted of words instead of digits. These words were drawn from the key word lists of McNally and Murray (1962) and Edwards and Gibbons (1964). No abstract nouns were used and no list contained more than one word of two syllables (the remainder all being words of one syllable). The words were randomised, but care was taken that no word appeared twice in any one list and no list contained word pairs likely to be meaningful to young children (e.g. 'baby cat', 'toy shop'). As in Test 1, the lists were presented to the children in ascending length order, and the test was terminated for each subject when (s)he had twice failed at a given length.

Test 3. Visual Digit Span. The lists of random digits used in this test were prepared as in Test 1, but were presented visually on a memory drum with one-second intervals between each digit. The signal for recall at the end of each list was a green circle. Scoring followed the same pattern as in Tests 1 and 2.

Test 4. Visual Word Span. The lists of words used in this test were prepared as in Test 2, but were presented visually on a memory drum as in Test 3. Scoring followed the pattern of the other tests.

Equivalence of stimuli presentation between auditory and visual modes was therefore secured as far as is practical by the use of identical stimulus material (words and digits) in each mode selected according to the same rationale, and by holding constant the non-experimental variables in the test procedures (e.g., length of stimulus lists, time interval between stimulus units).

Procedure

All subjects were tested individually. The auditory tests (Tests 1 and 2) were presented first over a two-week period, followed by the visual tests (Tests 3 and 4) also over two weeks. Since time of day appears to influence memory efficiency (Blake, 1967; Baddeley *et al.*, 1970) all testing was carried out during the first half of the school day between the hours of 9 a.m. and 12 noon.

Results were analysed in three ways, firstly for the sample as a whole, and then for each sex separately. In each of the three ways year groups were treated discretely. Tables 1 to 3 show mean differences and 't' statistics for the two modes of digital presentation analysed in these three ways, and Tables 4 to 6 display equivalent data for the word tests. In the tables the following abbreviations are used:

ADS—Auditory Digit Span

AWS—Auditory Word Span

VDS—Visual Digit Span

VWS—Visual Word Span

Table 1 indicates a very significant difference in mean scores for each of the year groups between the two modes of presentation. In each case it is the auditory mode which yields the higher mean score. This pattern is repeated when the sexes are treated separately (Tables 2 and 3).

TABLE 1

DIFFERENCES OF MEANS, AND t STATISTICS FOR AUDITORY AND VISUAL DIGITAL PRESENTATION BY YEAR GROUPS (WHOLE SAMPLE)

Year Group	N	Mean Age	ADS		VDS		Diff. of Means	t
			Mean	SD	Mean	SD		
1	19	7.7	5.00	0.75	3.80	0.79	1.21	7.40**
2	25	8.8	5.36	0.76	4.12	0.83	1.24	6.13**
3	28	9.7	6.04	1.04	4.25	0.84	1.79	8.58**
4	27	10.4	5.59	0.80	4.41	1.31	1.19	5.73**
Total	99	9.2	5.55	0.92	4.17	0.99	1.37	13.32**

Mean ages quoted in years and months. ** P < 0.01

TABLE 2

DIFFERENCES OF MEANS, AND t STATISTICS FOR AUDITORY AND VISUAL DIGITAL PRESENTATION BY YEAR GROUPS (BOYS ONLY)

Year Group	N	Mean Age	ADS		VDS		Diff. of Means	t
			Mean	SD	Mean	SD		
1	13	7.7	5.00	0.71	3.77	0.83	1.23	5.33**
2	14	8.8	5.29	0.73	4.14	0.94	1.14	4.16**
3	10	9.7	6.10	1.10	4.40	1.08	1.70	4.02**
4	15	10.5	5.27	0.70	4.20	1.08	1.07	4.30**
Total	52	9.1	5.37	0.86	4.12	0.98	1.25	8.78**

Mean ages quoted in years and months. ** P < 0.01

TABLE 3

DIFFERENCES OF MEANS, AND *t* STATISTICS FOR AUDITORY AND VISUAL DIGITAL PRESENTATION BY YEAR GROUPS (GIRLS ONLY)

Year Group	N	Mean Age	ADS		VDS		Diff. of Means	<i>t</i>
			Mean	SD	Mean	SD		
1	6	7.6	5.00	0.89	3.83	0.75	1.17	7.00**
2	11	8.8	5.46	0.22	4.09	0.70	1.36	4.40**
3	18	9.7	6.00	1.03	4.17	0.71	1.83	7.90**
4	12	10.4	6.00	0.74	4.67	1.58	1.33	3.75**
Total	47	9.3	5.75	0.94	4.23	1.00	1.51	10.17**

Mean ages quoted in years and months. ** $P < 0.01$

TABLE 4

DIFFERENCES OF MEANS, AND *t* STATISTICS FOR AUDITORY AND VISUAL WORD PRESENTATION BY YEAR GROUPS (WHOLE SAMPLE)

Year Group	N	Mean Age	AWS		VWS		Diff. of Means	<i>t</i>
			Mean	SD	Mean	SD		
1	19	7.7	4.37	0.76	3.42	0.77	0.95	6.65**
2	25	8.8	4.68	0.80	3.68	0.63	1.00	6.55**
3	28	9.7	5.36	0.78	4.07	0.72	1.29	10.33**
4	27	10.4	4.70	0.91	3.96	0.85	0.74	3.51**
Total	99	9.2	4.82	0.89	3.82	0.77	1.00	11.94**

Mean ages quoted in years and months. * $P < 0.01$

TABLE 5

DIFFERENCES OF MEANS, AND *t* STATISTICS FOR AUDITORY AND VISUAL WORD PRESENTATION BY YEAR GROUPS (BOYS ONLY)

Year Group	N	Mean Age	AWS		VWS		Diff. of Means	<i>t</i>
			Mean	SD	Mean	SD		
1	13	7.7	4.39	0.65	3.31	0.75	1.08	6.06**
2	14	8.8	4.50	0.94	3.64	0.63	0.86	3.71**
3	10	9.7	5.40	1.08	4.20	0.63	1.20	4.81**
4	15	10.5	4.53	0.83	3.67	0.49	0.86	4.52**
Total	52	9.1	4.65	0.93	3.67	0.68	0.98	9.38**

Mean ages quoted in years and months. ** $P < 0.01$

TABLE 6

DIFFERENCES OF MEANS, AND *t* STATISTICS FOR AUDITORY AND VISUAL WORD PRESENTATION BY YEAR GROUPS (GIRLS ONLY)

Year Group	N	Mean Age	AWS		VWS		Diff. of Means	<i>t</i>
			Mean	SD	Mean	SD		
1	6	7.6	4.33	1.03	3.67	0.82	0.67	3.16*
2	11	8.8	4.91	0.54	3.73	0.65	1.82	6.50**
3	18	9.7	5.33	0.59	4.00	0.77	1.33	9.52**
4	12	10.4	4.92	1.00	4.33	1.07	0.58	1.40
Total	47	9.3	5.00	0.81	3.98	0.85	1.02	7.61**

Mean ages quoted in years and months. * $P < 0.05$. ** $P < 0.01$.

When verbal presentation is analysed in the same three ways (Tables 4-6) we find closely similar results, with the auditory mode again producing consistently superior short-term memory recall to the visual mode. Only two out of the 15 obtained 't' statistics fail to reach significance at the 0.01 level (Table 6) though both of these are in the right direction and one attains significance at the 0.05 level.

DISCUSSION

Results provide strong support for the view that aurally presented digits and words prompt more effective short-term memory function than visually presented. So marked is the effect that it remains statistically highly significant even when the sample is broken down by age and sex into relatively small sub-groups. Interestingly, although memory span for both digits and words in both modes of presentation tends to increase in the present sample with age, there is no evidence that the superiority of the aural mode is itself a function of age. No clear pattern emerges across the year groups, and one is driven to conclude that at all stages of junior schooling material presented aurally is likely to evoke superior short-term memory performance in children than material presented visually (at least where this material is in the form of digits and words). This conclusion holds equally for both sexes.

Without doubt, therefore, our results support those of Dornbush and Basow (1970) and Baumeister (1974) who found aural presentation to be superior to visual, and go against those of Siegal and Allik (1973), Perlmutter and Myers (1975), Spitzer (1976) and Koppitz (1977) who found to the contrary. Present findings also emphasise the importance of test procedures that strive for equivalence of stimuli presentation and hold constant non-experimental variables. It was probably failure to do this that accounts for the results obtained by Koppitz and other workers who found visual presentation apparently preferable to aural. The use of equivalent digital and verbal material in both modes also leaves us better able to enter theoretical speculation on the mechanisms of short-term memory, since such equivalence confronted subjects with comparable encoding tasks in these modes. Results support the model of short-term processing advanced by Sperling (1963), namely that visual information must be coded acoustically if it is to be lodged in the short-term store, and that in consequence aural material, which can be coded in the form in which it is received, has obvious advantages over visual in maximising short-term span efficiency.

Given that we could generalise from the above findings, there would be clear practical implications for education. Reference has already been made to the emphasis placed in much of primary education on visual as opposed to aural experience, but reference should also be made to student study habits at all age levels. The potential superiority of tape-recorded material played back to students over printed material perused from notes or textbooks is an obvious focus for further research, and the design and presentation of audio-visual aids and of the devices utilised by educational technology might also require to be rescrutinised if the findings of the present study are confirmed in future more wide-ranging investigations.

REFERENCES

- BADDELEY, A. D., HATTER, J. E., SCOTT, D., and SNASHALL, A. (1970). Memory and time of day. *Br. J. Psychol.*, 22, 605-609.
- BASU, A. K. (1972). An experimental study in memory span. *Indian J. Exper. Psychol.*, 6, 64-66.
- BAUMEISTER, A. A. (1963). *Investigation of Memory Deficits in Retardation*. Progress Report MHO 7445-01, National Institute of Mental Health (USA).
- BAUMEISTER, A. A. (1974). Several memory span thresholds of normal and mentally retarded children. *J. educ. Psychol.*, 66, 889-894.
- BLAKE, M. J. F. (1967). Time of day effects in a range of tasks. *Psychon. Sci.*, 9, 349-350.
- DORNBUSH, R. L., and BASOW, S. (1970). The relationship between auditory and visual short term memory and reading achievement. *Child Dev.*, 41, 1033-1044.

- EDWARDS, R. P. A., and GIBBONS, U. (1964). *Words Your Children Use: A graded Vocabulary*. London: Burk.
- GUTHRIE, J. T., and GOLDBERG, H. K. (1972). Visual sequential memory in reading disability. *J. Learning Disab.*, 5, 41-46.
- KATZ, P. A., and DEUTSCH, M. (1967). Auditory and visual functioning and reading achievement. In DEUTSCH, M. (Ed.), *The Disadvantaged Child*. New York: Basic Books.
- KOPFITZ, E. M. (1977). *The Visual Aural Digit Span Test*. New York: Grune.
- LUNZER, E. A., DOLAN, T., and WILKINSON, J. E. (1976). The effectiveness of measures of operativity, language and short-term memory in the prediction of reading and mathematical understanding. *Br. J. educ. Psychol.*, 46, 295-305.
- M McNALLY, J., and MURRAY, W. (1962). *Key Words to Literacy and the Teaching of Reading*. London: Schoolmaster Publishing Co.
- MURRAY, D. J., and ROBERTS, B. (1968). Visual and auditory presentation rate and short-term memory in children. *Br. J. Psychol.*, 59, 119-125.
- PERLMUTTER, M. and MYERS, N. A. (1975). Young children's wording and storage of visual and verbal material. *Child Dev.*, 46, 215-219.
- SIEGAL, A. W. and ALLIK, J. P. (1973). A developmental study of visual and auditory short term memory. *J. verb. Learn. verb. Behav.*, 12, 409-418.
- SPEERLING, G. (1963). A model for visual memory tasks. *Human Factors*, 5, 19-31.
- SPITZER, T. M. (1976). The development of visual and auditory recall as a function of presentation and probe modalities, serial position, and series size. *Child Dev.* 47, 767-778.
- TORGESON, J. K. (1975). Memorization processes in reading disabled children, *J. Educ. Psychol.*, 65, 571-578.

(Manuscript received 4th February, 1980)

APPENDIX

TEST 1

List of words used in auditory span test. The third list of each series length was included for use in the event of either of the first two of that series length being invalidated (e.g., by tape recorder malfunction).

Practice Lists father cup; boy farm woman

Actual Test

- | | | | |
|-----|--|-----|--|
| 2 { | boy tea
apple girl
pen apple | 5 { | shop rabbit book home hill
train dog hill road man
sea baby train street toy |
| 3 { | train horse car
pen water bus
flower milk queen | 6 { | rain bus cap queen father tea
girl toy sun dog tea shop
rabbit fire toy rain head jam |
| 4 { | box pig table head
van fire hand money
dinner cap boat chair | 7 { | water top up ball street door book
train dog cap window hat school tea
hill van cat letter dog box van |
| 8 { | cat house man hat bus cup jam street
train picture sun head street book game hill
bus sister van egg hill cake toy road | | |
| 9 { | fish table man door bag road cake horse milk
window book boat bird seat bus ball pig hat
door cake cat baby toy sea pig car boat | | |

TEST 2

List of digits used in auditory span test. The third list of each series length was included for use in the event of either of the first two of that series length being invalidated.

Practice Lists 8 5; 2 4 9

Actual Test

- | | | | |
|-----|-------------------|-----|-------------------------|
| 2 { | 3 1
6 8
9 2 | 3 { | 2 9 4
7 3 4
5 9 1 |
|-----|-------------------|-----|-------------------------|

$$4 \begin{cases} 9 & 3 & 2 & 7 \\ 8 & 3 & 4 & 2 \\ 1 & 6 & 9 & 4 \end{cases}$$

$$6 \begin{cases} 4 & 1 & 3 & 7 & 2 & 5 \\ 3 & 6 & 7 & 4 & 8 & 2 \\ 6 & 3 & 8 & 4 & 5 & 1 \end{cases}$$

$$8 \begin{cases} 9 & 4 & 3 & 8 & 5 & 6 & 1 & 2 \\ 1 & 7 & 2 & 6 & 8 & 9 & 5 & 3 \\ 3 & 4 & 2 & 7 & 1 & 8 & 6 & 5 \end{cases}$$

$$5 \begin{cases} 3 & 2 & 8 & 1 & 5 \\ 4 & 1 & 6 & 5 & 8 \\ 4 & 9 & 2 & 1 & 7 \end{cases}$$

$$7 \begin{cases} 6 & 4 & 1 & 5 & 7 & 2 & 9 \\ 4 & 5 & 9 & 2 & 3 & 6 & 7 \\ 2 & 4 & 1 & 3 & 6 & 8 & 5 \end{cases}$$

$$9 \begin{cases} 7 & 3 & 1 & 2 & 5 & 6 & 4 & 8 & 9 \\ 8 & 4 & 6 & 5 & 7 & 1 & 3 & 2 & 9 \\ 7 & 3 & 1 & 9 & 2 & 4 & 6 & 5 & 8 \end{cases}$$

TEST 3

List of words used in visual span test. The third list of each series length was included for use in the event of either of the first two of that series length being invalidated (e.g., by memory drum malfunction).

Practice Lists cake fire; dinner van pen;

Actual Test

$$2 \begin{cases} \text{toy head} \\ \text{chair ball} \\ \text{mother hand} \end{cases}$$

$$3 \begin{cases} \text{cup rain top} \\ \text{woman hat house} \\ \text{letter bird rain} \end{cases}$$

$$4 \begin{cases} \text{picture dog book car} \\ \text{milk head rain woman} \\ \text{man game bed station} \end{cases}$$

$$5 \begin{cases} \text{man book van water door} \\ \text{flower home book game shop} \\ \text{money pen milk room van} \end{cases}$$

$$6 \begin{cases} \text{rabbit sea cake farm toy fire} \\ \text{egg road door rain picture van} \\ \text{bird picture sun boat jam fish} \end{cases}$$

$$7 \begin{cases} \text{horse pen water bird car box milk} \\ \text{train bed shop horse farm picture tea} \\ \text{milk road cat top game window man} \end{cases}$$

$$8 \begin{cases} \text{egg ball fish woman house sea brain bird} \\ \text{water car sea milk egg ball rain van} \\ \text{father shop pig doll van hat milk bus} \end{cases}$$

$$9 \begin{cases} \text{shop road cat top cow room bed hand fire} \\ \text{box window pen bus farm milk dog hand cake} \\ \text{sister book home tea fire boy cat bag toy} \end{cases}$$

TEST 4

List of digits used in visual span test. The third list of each series length was included for use in the event of either of the first two of that series length being invalidated.

Practice Lists 5 2; 8 4 1

Actual Test

$$2 \begin{cases} 8 & 6 \\ 3 & 8 \\ 6 & 2 \end{cases}$$

$$3 \begin{cases} 9 & 8 & 3 \\ 5 & 3 & 6 \\ 4 & 7 & 1 \end{cases}$$

$$4 \begin{cases} 5 & 2 & 9 & 8 \\ 9 & 2 & 5 & 7 \\ 3 & 8 & 6 & 5 \end{cases}$$

$$5 \begin{cases} 3 & 2 & 7 & 5 & 6 \\ 4 & 9 & 1 & 2 & 7 \\ 9 & 6 & 4 & 8 & 1 \end{cases}$$

$$6 \begin{cases} 1 & 6 & 5 & 2 & 4 & 8 \\ 5 & 9 & 2 & 8 & 7 & 3 \\ 7 & 5 & 8 & 1 & 6 & 2 \end{cases}$$

$$7 \begin{cases} 1 & 3 & 4 & 2 & 7 & 9 & 6 \\ 5 & 3 & 1 & 8 & 5 & 7 & 6 \\ 9 & 4 & 8 & 1 & 5 & 2 & 7 \end{cases}$$

$$8 \begin{cases} 5 & 8 & 7 & 2 & 1 & 6 & 4 & 9 \\ 4 & 2 & 5 & 3 & 8 & 9 & 6 & 7 \\ 2 & 9 & 1 & 6 & 4 & 8 & 7 & 5 \end{cases}$$

$$9 \begin{cases} 6 & 2 & 9 & 5 & 3 & 8 & 4 & 1 & 7 \\ 5 & 7 & 1 & 9 & 4 & 3 & 6 & 2 & 8 \\ 3 & 8 & 6 & 1 & 5 & 2 & 9 & 7 & 4 \end{cases}$$

THE EFFECTS OF A DYSLEXIA-CENTRED TEACHING PROGRAMME

By BEVÉ HORNSBY* and T. R. MILES

(Department of Psychology, University College of North Wales, Bangor)

SUMMARY. Results are presented for 107 retarded children who received tuition by means of a 'dyslexia-centred' programme. In 41 cases the teaching was organised by a hospital clinic, in 44 cases by a unit attached to a university department, and in 22 cases by a private body, these three centres being different in many other ways but alike in that all took the concept of dyslexia seriously. It is argued that in suitable conditions such teaching can be effective, and the implications for the concept of dyslexia are briefly discussed.

INTRODUCTION

THIS paper gives the results, in terms of reading age and spelling age, for 107 retarded children who were taught by methods which can be described as 'dyslexia-centred'. Records were obtained from three different centres: Centre A was a clinic attached to a hospital; Centre B was a unit attached to a university department, while Centre C was organised on a private basis. All the children referred to had been diagnosed as dyslexic.

The clinical pattern of dyslexia, which occurs with only relatively minor variations, has been described by many authors (cf Miles, 1978): it involves not merely lateness in learning to read and retardation at spelling but other distinctive difficulties, for example in working out left and right, in reciting tables and other items which occur in sequence, and in the recall of visually and auditorily presented digits. For a variety of reasons the source of these difficulties is believed to be constitutional, and it seems likely that it involves some sort of limitation in immediate memory for verbal material (cf Ellis and Miles, 1981; Miles and Ellis, 1981).

At Centre A the book which formed the basis of the teaching programme was Hornsby and Shear (1977); at Centre B a syllabus was used which involved extensions and adaptations of Miles (1970) and Miles and Miles (1975), while at Centre C the teaching was based primarily on Hickey (1977). Also in use were the Gill Cotterill checklist and the Edith Norrie letter case. The pioneer work of Gillingham and Stillman (latest edition, 1969) has of course had considerable influence on the teaching of dyslexic children in many parts of the world. (For a recent survey of dyslexia-centred methods see Naidoo, 1981.)

If the above books are studied in detail the basic similarities will soon become apparent. Rawson (1970) has said that the teaching programme for a dyslexic pupil needs to be 'structured, sequential, cumulative, and thorough'. All the books in their various ways satisfy these criteria. A *structured* programme is one which has a specific organisation: the different parts cohere and contribute to a planned whole. If the teaching is carried out *sequentially* this means that there is a progressive disclosure by the teacher of the correspondence between sounds and letters (or groups of letters) in the English spelling system. For example, the pupil is taught at an early stage that every word contains a vowel; somewhat later he is taught that the vowel sound is different according to whether or not there is a final 'e' (e.g., *hat*, *hate*) and that two vowels together give a long vowel sound (e.g., *sleep*, *boat*). (Some of these points, e.g., that relating to the final 'e', are of course taught quite independently of the concept of dyslexia, but when this happens one sometimes receives the impression that the children are being shown 'bits and pieces' rather than working through a coherent programme with a determinate structure.) Thirdly, the teaching needs to be

* Now at the Dyslexia Clinic, St Bartholomew's Hospital, London.

cumulative, in the sense that each stage follows from what went before and cannot be studied in isolation. For example there would be no point in studying consonant blends before the pupil knows the sounds made by individual consonants and there would be no point in teaching the prefix 'dis' before the pupil was familiar with the short i-sound in *hit*, *bit*, etc. Finally the teaching needs to be *thorough* in the sense that each stage in the programme must be fully understood before the pupil passes to the next one.

It is important to emphasise that the training given at the three centres involved not simply 'remedial reading' but the teaching of a range of language skills of a more diverse kind.

As far as the authors could tell, differences in approach between the three centres were minimal. There was some disagreement as to whether the *names* of the letters should be taught from the start: the organiser of Centre B regarded this as unnecessary; the organiser of Centre A thought it desirable, while the organiser of Centre C thought it essential. The authors received the impression that the methods used by all three centres were of the 'brick-by-brick' variety but that there was somewhat less rigidity in this respect at Centre B as compared with Centre C, with Centre A occupying an intermediate position. It was not possible to check on differences in approach between teachers at the same centre. They were satisfied, however, that all three centres had common aims and that they were following what was basically the same programme.

Tuition was free in the case of Centres A and B, while children attending Centre C were paid for either privately or as a result of help from LEAs. In the case of Centres A and B the children were taught on an individual basis, except in a very few cases when the Centre B children were taught in pairs. In the case of Centre C up to four children were taught together, though each child had his own individual programme. In the case of Centres A and C the children attended at the centre for their lessons; in the case of Centre B teaching was carried out in the school by teachers visiting on a peripatetic basis, assessment having previously been carried out by one of the county educational psychologists. At Centre C the norm was two lessons per week; at Centres A and B it was one lesson per week, though in the case of Centre B a very small number of pupils received two lessons per week for short periods during the total time of tuition.

PROCEDURE

Tests of reading and spelling had been given at all three centres, in most cases the Neale Analysis of Reading Ability, the Schonell R_1 word-recognition test and the Schonell S_1 spelling test. In the case of the Neale test the 'accuracy' score was used, and R_1 scores were converted, where necessary, so as to reflect the more recent norms. To ensure a standardised procedure only those children were included whose IQ had been assessed as 95 or more and whose spelling age was less than 80 per cent of their chronological age; the original list of 183 names was thus reduced to 107. Of the 85 children whose sex was given 71 were boys. (For evidence of sex differences in dyslexia see Critchley, 1970.) According to the records 51 of the children were within the IQ range 95-109, 36 within the range 110-119, and 20 with a figure of 120 or above.

There were both ethical and practical difficulties in trying to obtain a matched control group of untreated pupils, and it seemed more appropriate to use each subject as his own control and to compare rate of gain *before* teaching with rate of gain *during* teaching. In accordance with established practice it was assumed that a child's reading and spelling education starts at the age of 5 years, and the before-teaching rate of gain was calculated by subtracting 60 months from the before-teaching reading age and dividing the result by the before-teaching chronological age less 60 months. Thus if a child on being tested before teaching was aged 10 years 3 months and had a reading

age of 6 years 9 months, his reading gain would be 21 months in 5 years 3 months (i.e., 63 months) and the rate of reading gain during this period would therefore be 21/63 or 0.33. A rate of gain of more than unity thus indicates that the child was 'keeping up with the clock'. Rate of spelling gain was calculated in the same way. In the case of both reading and spelling the difference between during-teaching rate of gain and pre-teaching rate of gain is taken to be an indication of the effectiveness of the teaching. The null hypothesis ('that teaching is ineffective') predicts no difference between the two.

The testing procedures and record keeping were not identical at the three centres. Pupils at Centre A were tested routinely whether or not they were ready for immediate discharge, and the records supplied by this centre related to all children on their books who were tested during the summer of 1977; 'time taught' thus means 'time between first and second testing'. In contrast Centres B and C tested the children at the start of tuition and immediately before discharge, and in their case 'time taught' means total time taught before discharge. The records from Centre B related to all those discharged during a three-year period (July, 1974 to December, 1977), with the exception, so the authors were told, of a small number of pupils whose teaching was discontinued for extraneous reasons, for example because the family left the district. The records for Centre C related to all children who had been taught for longer than six months during the period January, 1973 to July, 1975.

It was, of course, necessary to ensure that children had not been included in the records precisely because they had been successful. The existence of routine testing was a sufficient safeguard in the case of Centre A, while in the case of Centres B and C the records relate to all those who were discharged over a specified period of time (41 months and 26 months respectively); and it follows that if a child from either centre was not ready for discharge at a given point in time his name would be carried forward into the next year and not simply omitted.

As a result of these precautions it can reasonably be claimed that the data relate to an unbiased sample of dyslexic children (or at the very least a sample of spelling-retarded children) and that generalisation to a wider sample of similar children is legitimate.

RESULTS

On the basis of the data supplied by the three centres in respect of the pupils selected for inclusion a chart was prepared which contained the following information: chronological age, reading age, and spelling age on first testing, chronological age, reading age and spelling age on second testing, and (consequently) time taught between the two tests, reading gain, and spelling gain. Six further columns were then calculated, viz. rate of reading gain before teaching (RB), rate of spelling gain before teaching (SB) (both calculated by the method described in section 2), rate of reading gain during teaching (RD) (calculated by dividing reading gain by time taught), rate of spelling gain during teaching (SD) (calculated by dividing spelling gain by time taught), and finally an RD-RB column by which the two rates of reading gain were compared and an SD-SB column by which the two rates of spelling gain were compared. The null hypothesis ('that teaching is ineffective') predicts zero values for both $\Sigma(RD-RB)$ and $\Sigma(SD-SB)$.

The relevant information can be summarised in Tables 1 and 2:

The figures in Table 2 represent the gains in reading age and spelling age per unit time, e.g., gain in months per month or years per year. A zero or negative score for (RD-RB) occurred in the case of six pupils and a zero or negative score for (SD-SB) in the case of three pupils.

TABLE 1
BASIC DATA

	Mean	SD	No. of children
Age at start of teaching (months)	127.39	25.08	107
Time between 1st and 2nd testing (months)	17.50	12.42	107
Duration of teaching (months) (Centres B and C)	20.29	9.61	66
Reading age at start of teaching (months)	94.66	16.69	107
Spelling age at start of teaching (months)	82.08	14.46	107
Reading age gain during teaching (months)	27.68	16.24	102
Spelling age gain during teaching (months)	27.67	14.96	107
Reading age rate of gain before teaching (years per year)	0.53	0.19	107
Spelling age rate of gain before teaching (years per year)	0.32	0.18	107
Reading age rate of gain during teaching (years per year)	1.91	1.11	102
Spelling age rate of gain during teaching (years per year)	1.94	1.07	107

TABLE 2
COMPARISON OF BEFORE-TEACHING RATES OF GAIN AND DURING-TEACHING RATES OF GAIN

	Mean	SD
Reading (RD-RB)	1.40	1.10 ($t = 12.8$, $df\ 101$, $P < 0.001$)
Spelling (SD-SB)	1.62	1.10 ($t = 15.0$, $df\ 106$, $P < 0.001$)

A number of different criteria for 'success' seem possible. The first and obvious one is that of whether the child 'kept up with the clock' during the period of teaching. It follows, of course, from the method of selection that at the start of teaching no child had 'kept up with the clock' in the case of spelling, and in fact only two children out of the 107 had 'kept up with the clock' in the case of reading, although nine others had a pre-teaching rate of reading gain of 0.8 or 0.9 and were thus not far retarded. During teaching, however, 91 children (85 per cent) had a reading gain of more than unity and a further eight (7 per cent) had a reading gain of more than 0.8. In the case of spelling there were again 91 children (85 per cent) who had a gain of more than unity and a further eight (7 per cent) who had a gain of more than 0.8. Only five children (5 per cent) achieved a rate of gain of under 0.8 at both reading and spelling.

It may be preferable, however, to set more ambitious criteria for success. This can be done if one thinks in terms of targets to be aimed at. What is an appropriate target may vary between one case and another for all kinds of reasons, but if it is asked what is a realistic target for a child of average ability by the time he reaches secondary school, an approximate answer might be 'reading age 10.0, spelling age 9.0'. If these levels are then assumed to provide reasonable targets for pupils in general, success rate can be measured by determining what percentage of pupils achieved them. An inspection was therefore made of the figures for Centres B and C (no figures being available for Centre A because of their policy on re-testing, see above) and it was found that 44 pupils out of 66 (67 per cent) achieved the reading target and 42 pupils out of 66 (64 per cent) achieved the spelling target. A further 11 pupils (17 per cent) achieved a reading age of 9.6 and a further 12 pupils (18 per cent) achieved a spelling age of 8.6; thus by a less stringent criterion the success rates were 83 per cent for reading and 82 per cent for spelling. Success and failure rates can be summarised as follows:

TABLE 3
FAILURE RATES BY DIFFERENT CRITERIA

	Reading		Spelling	
	% success	% fail	% success	% fail
Kept up with the clock	85	15	85	15
0.8 of keeping up with the clock	93	7	93	7
Achieved reading age of 10.0	67	33	—	—
Achieved reading age of 9.6	83	17	—	—
Achieved spelling age of 9.0	—	—	64	36
Achieved spelling age of 8.6	—	—	80	20

Comparison between one centre and another seemed inappropriate. Not only would such comparison be somewhat invidious, but, more important, it might be positively misleading: in particular, a lower place in the 'league table' might be due not to less efficient teaching but simply to a willingness to take on less promising cases. The authors satisfied themselves, however, that all three centres were contributing to the success rate: thus it was found, for instance, that by the criterion of 0.8 of 'keeping up with the clock' no centre had a failure rate of over 9 per cent at either reading or spelling. Spectacular improvements were found at all three centres.

DISCUSSION AND CONCLUSIONS

The basic facts speak for themselves. The null hypothesis ('that teaching is ineffective') was refuted in the case of both reading and spelling at a level of confidence which leaves no room for doubt. Reading gains in the during-teaching period averaged 1.91 years per year and spelling gains 1.94 years per year, compared with averages in the pre-teaching period of 0.53 and 0.32 respectively.

With regard to failure rate, it does not seem appropriate to lay down criteria which could be adopted universally since conditions may vary at different teaching centres. On the basis of our results, however, it seems safe to say that in the case of dyslexic children a high failure rate is unnecessary.

The reasons for the success, however, require further discussion. It is sometimes said that any teaching method will work provided the teacher can study the child's individual needs, is kind to him, and gives him confidence. But is one to assume that before the child was given dyslexia-centred teaching no one else tried to study his needs, was kind to him, or tried to give him confidence? Although records of what special teaching the children received before they had dyslexia-centred teaching are not available it is worth while considering the circumstances in which the referrals took place. The organiser of Centre A writes: "All the children in our sample had already received remedial help in school before they came to the clinic and had been seen by the LEA educational psychologist regarding either reading failure or behaviour problems". Moreover, since assessment at Centre A was possible only if referral was made through a doctor, one must assume that the problem seemed to the parents to be sufficiently pressing to justify the seeking of help through medical channels, which scarcely reflects confidence in what was being done educationally. Although at all three centres the 'remedial help in school' may have varied in quality and extent between one child and another, it is hard to believe that total indifference is the norm. Indeed the experience of the authors suggests that the conscientious class teacher is often bewildered by the dyslexic child and does not know what his needs are or how to begin the task of giving him confidence. In these circumstances kindness does not go far, and if one tells a child that he *ought* to have more confidence in himself without at the same time showing him how to achieve success, one is merely adding to his diffi-

culties by adverse criticism. It is important, of course, not to underestimate the influence of the so-called Hawthorn effect (see, e.g., Sprott, 1952, pp. 234-237), i.e., the improvement in morale resulting from the realisation that someone is genuinely trying to help, regardless of how the help is given. But most of these children, especially those in groups A and C, had come from a caring environment: their families were willing actively to seek help from a hospital clinic or a private agency because their own attempts at kindness had not helped their child's reading or spelling. In these conditions it is hard to believe that the Hawthorn effect on its own was the decisive factor. On the contrary, the most likely explanation of the results is the straightforward one, viz., that the reading and spelling of dyslexic children can be improved if the right kind of teaching programme is used.

It is also possible to make an appeal to coherence. The methods described earlier in this paper did not evolve simply by accident; it is not that they just *happen* to be effective with dyslexic children. On the contrary, it is known that dyslexic children are strong at reasoning and weak at tasks which impose any appreciable load on immediate memory, and it is therefore only to be expected that a method which encourages deduction and discourages rote learning will be more successful than one, like 'look-and-say', which calls for the absorption of a large amount of detail in a short time (cf E. Miles, 1981).

The data do not, of course, establish that no other programme would have been successful, but the onus now rests with those whose programmes are different and/or not dyslexia-centred to demonstrate their effectiveness.

It may still be objected, however, that the programme which has been described in this paper is one which should be used by *any* teacher whether or not he accepts the concept of dyslexia, and it does not follow in that case that its success justifies the classification of certain children as 'dyslexic'. Now certainly the important thing is that any effective programme should be widely used, and, indeed, little is lost if a programme devised for those with a reduced span of immediate memory is also used with those whose immediate memory is normal. The converse, however, does not hold: if there is a reduced span of immediate memory, only certain procedures will be effective while others will be not only a waste of time but positively discouraging.

However that may be, the evidence in this paper certainly shows that even children with seemingly obstinate reading and spelling problems can nevertheless be helped; and it follows that considerably more effort is needed on a nation-wide scale to ensure that such help is given.

REFERENCES

- CRITCHLEY, M. (1970). *The Dyslexic Child*. London: Heinemann Medical Books.
- ELLIS, N. C., and MILES, T. R. (1981). A lexical encoding deficiency, I. In PAVLIDIS, G. TH., and MILES, T. R. (Eds.), *Dyslexia Research and Its Application to Education*. Chichester: John Wiley (in press).
- GILLINGHAM, A., and STILLMAN, B. E. (1969). *Remedial Training for Children with Specific Disability in Reading, Spelling and Penmanship*. Cambridge, Mass.: Educators Publishing Service.
- HICKEY, K. (1977). *Dyslexia. A Language Training Course for Teachers and Learners*. 3 Montague Road, London SW19.
- HORNSBY, B., and SHEAR, F. (1977). *Alpha to Omega*. London: Heinemann Educational.
- MILES, E. (1981). A study of dyslexic weaknesses and the consequences for teaching. In PAVLIDIS, G. TH., and MILES, T. R. (Eds.), *Dyslexia Research and Its Application to Education*. Chichester: John Wiley (in press).
- MILES, T. R. (1970). *On Helping the Dyslexic Child*. London: Methuen Educational.
- MILES, T. R. (1978). *Understanding Dyslexia*. London: Hodder and Stoughton.
- MILES, T. R., and ELLIS, N. C. (1981). Dyslexia as a lexical encoding deficiency, II. In PAVLIDIS, G. TH., and MILES, T. R. (Eds.), *Dyslexia Research and Its Application to Education*. Chichester: John Wiley (in press).
- MILES, T. R., and MILES, E. (1975). *More Help for Dyslexic Children*. London: Methuen Educational.

THE DEVELOPMENT OF SEX DIFFERENCES IN EVALUATIONS OF OTHERS' SUCCESSES AND FAILURES

By C. G. ROGERS

(*Department of Educational Research, University of Lancaster*)

SUMMARY. 240 children aged 9, 12 and 15 years evaluated successes and failures of other children. Information concerning the other children's levels of ability, effort and outcome together with sex was supplied. A number of sex differences were found particularly concerning the extent to which the evaluations related to the effort and outcome information supplied. The results are compared to those of earlier studies conducted in the United States of America and Iran.

INTRODUCTION

THE last five years or so have seen a considerable amount of research activity associated with Weiner's (1974) attributional theory of achievement motivation. More recently attention has turned to developmental aspects of this particular approach, but, as shown in a review of developmental studies (Weiner and Kun, 1976), there are a considerable number of outstanding problems to be addressed.

The attributional theory of achievement motivation is essentially an elaboration of the familiar expectancy theory approach (Atkinson and Raynor, 1974). An individual's level of achievement motivation is in any one instance a function of, amongst other factors, his or her expectations of success. Weiner's contribution here has been to demonstrate how expectations for success can be mediated by the causal attributions made to explain past successes and failures. Attributions to stable causal factors (e.g., ability, task difficulty) lead to the expectation that future performance will closely match past ones. Attributions to unstable causal factors (e.g., effort, luck) reduce the strength of such a link. Further, individuals develop attributional styles leading them typically to make similar attributions over different events with respect to success and failure (Ickes and Layden, 1978). The making of such causal attributions is known to be a complex process. Information about the presence or absence of possible causal factors has to be noticed, evaluated, selected and integrated.

Little is known about the development of any of these sub-processes. The problem of how a child learns to recognise the cues associated with various causal factors has been virtually ignored—largely, one suspects due to the incompatibility of such a question with the accepted experimental 'paradigm' of attribution research. Integration of such information has received more attention largely within the context of Kelley's (1972) ideas on the nature of causal schemata. For example, Smith (1975) and Schultz *et al.*, (1975) have shown that Kelley's multiple sufficient causal schema is not present until after the age of 7-9 years. The present paper is concerned with developmental changes in the evaluation of causal factors associated with achievement.

Weiner and Peter (1973) examined children's evaluations of the causal factors effort and ability together with their evaluations of actual outcomes of achievement activities. They also compared the development of achievement judgments with moral judgments but that is not of present concern.

Children aged 4 to 18 years rewarded or punished others described as having either succeeded or failed at a puzzle completion task with all possible combinations of high or low effort and high or low ability. The results showed that success is rewarded more than failure, high effort more than low effort and that amongst the oldest children lack of ability more than its presence. This last finding reflects the high

rewards given by these children to those who succeeded in spite of a lack of ability together with the stern view taken of those with high ability who still failed.

More importantly, until the age of 10-12 years outcome was the principle determinant of level of reward. Effort increased in importance until at 10-12 years it became the most important of the determinants of reward level studied. After this age, however, a 'regression effect' was found with outcome again becoming dominant although effort continued to be seen as important. Weiner and Peter suggest that this developmental trend is the result of a combination of cognitive development and cultural influences. Young children are held to be relatively incapable of comprehending the concept of effort. Their evaluations are more readily made on the more concrete basis of outcome information. As children's (largely unspecified) cognitive skills develop so they are better equipped and more disposed to use the more subjective information about level of effort as the basis of their evaluations. However, they also come to learn that their culture places a greater value on products than it does on effort—hence the short-lived ascendancy of effort information as a determinant of evaluations of success and failure. In view of the well-established positive relationship between attributions for failure to lack of effort and adaptive achievement behaviour (e.g., Weiner and Kukla, 1970) such a developmental trend could be of considerable significance.

Salili *et al.* (1976) replicated Weiner and Peter's study in Iran. Their results were broadly comparable with Weiner and Peter's with two notable exceptions with respect to achievement evaluations. First, Iranian subjects' evaluations significantly varied as a function of ability whereas those in the US had not. Salili *et al.* state that: "Competence, independent of what one does with it, seems to be an inherent good in Iran—but not in the US" (p. 335). Second, Salili *et al.* did not find the age-related regression that was the prominent feature of Weiner and Peter's results. For the Iranian subjects effort was the most important determinant of reward level, and further this factor became steadily more important with age. Salili *et al.* conclude that their results support Weiner and Peter's claim that achievement evaluations are to be seen as the result of both cognitive development and cultural factors.

The present study seeks to replicate the study of Weiner and Peter in an English context with an additional factor taken into consideration. Neither Weiner and Peter nor Salili report a detailed analysis of sex differences in the evaluations made by their subjects and neither study attempts to examine systematically the sex of the person being evaluated. Weiner and Peter's results do include two significant interactions involving the sex of the judge but these are not discussed. Salili concludes that in spite of the greater distinction made between male and female roles in Iranian culture the sex of the judge did not affect the evaluations made.

There are, however, good reasons to believe that such differences might exist. Research reviewed by Crandall (1969) and Dweck (1979) shows consistent differences between boys and girls in terms of both their expectations for future success and failure and their attributions for past and present successes and failures.

Crandall reports a series of studies showing that girls will tend to underestimate their future performances while boys will tend to overestimate them. This applies in areas where the actual performance of the girls is as good as if not better than that obtained by the boys. Crandall also discusses data showing that these different expectations are not related to need for social approval or to 'masculinity-femininity' as assessed by the Californian Personality Inventory.

Research more directly related to attributional theory shows related sex differences. Dweck and Reppucci (1973) and Nicholls (1975) show that girls are more likely than boys to neglect lack of effort as a cause of failure and to emphasise instead a lack of ability. In accordance with predictions derived from attributional theory girls are

consequently less likely than boys to persist in the face of failure (Dweck and Gilliard, 1975; Nicholls, 1975). The research briefly outlined here indicates that perceptions of the relative importance of effort are central in determining typical sex differences in response to success and failure. The Weiner and Peter regression effect also suggests that cultural factors can influence these perceptions. A systematic investigation of judgments of success and failure as a function of both age and sex within a given cultural context is clearly called for.

METHOD

Sample

Subjects were recruited from primary and secondary schools, in the North-West of England. Two comprehensive secondary schools were used together with four primary schools which were all feeder schools for the two comprehensives. All schools were co-educational. Subjects were randomly drawn from unstreamed classes. The final sample contained 240 children, 80 from each of three age groups with half being boys. Average ages of the three groups were: 9 years 5 months, 12 years 1 month, and 15 years 9 months. These age groups correspond closely to those used by Weiner and Peter (who used groups aged 4-6, 7-9, 10-12, 13-15, and 16-18) and Salili (7-9 11-12, 13-15 and 16-18). It was in the age range included here that the most interesting results were obtained in the earlier studies, particularly those concerning the presence or absence of the regression effect.

Materials

A total of 16 versions of the following description were prepared and printed onto separate sheets:

"Bill is good at arithmetic. He tries hard to do the 10 problems. He gets 8 correct."

The variants were produced by changing the sex of the person described (by changing the name and pronoun); ability level (good or not good at arithmetic); effort level (tries hard or does not try very hard) and outcome (eight or two of the 10 problems correct). Beneath each description were printed the two questions associated with the rating system to be described below.

Procedures

Testing was conducted with groups of 15-20 children per session. Each child was given a booklet containing eight different versions of the descriptions. Children either made their judgments about boys only or girls only—the eight versions therefore being made up of all possible combinations of the outcome, effort and ability factors ($2 \times 2 \times 2$). Subjects were told that the descriptions gave accounts of how other children 'like them' had performed in a recent test. They were told that these children came from a school 'like theirs' but that it was one that they did not know and that they would not therefore know any of the children described to them. They were told to read and consider carefully all details of each description and then to evaluate each child. The procedure used here to record these evaluations was the same as that employed by Weiner and Peter. Gold stars were to be awarded to indicate that the subject was pleased with the other child and red to indicate displeasure. Up to five stars of either colour could be awarded producing a 10-point rating scale (scored from -5 to +5 with no zero). Children first responded to the question 'What colour star would you give him/her?' by encircling one of the two alternatives, and then to the question 'How many stars would you give him/her?' by encircling one of the numbers 1 to 5. It was explained to the two older groups that younger children were answering the same questions and that it was important that all responses were made in the same way.

In summary, the study employed a $3 \times 2 \times 2 \times 2 \times 2 \times 2$ design (age of subject $\times$ sex of subject (SexS) $\times$ sex of person judged (SexO) $\times$ ability $\times$ effort $\times$ outcome) with repeated measures over the last three factors. Each cell in the design contained a total of 20 responses.

RESULTS

Data from the study were subjected to a six-way analysis of variance with repeated measures over the last three factors. All main effects and all significant interactions derived from this analysis are shown in Table 1.

TABLE 1

SUMMARY OF ANALYSIS OF VARIANCE SHOWING ALL MAIN EFFECTS AND SIGNIFICANT INTERACTIONS

Source	df.	ms.	F.	Signif. level
Age	2	4.51	0.86	NS
SexS	1	2.70	0.51	NS
SexO	1	0.41	0.08	NS
Ability	1	6.77	1.84	NS
Effort	1	2108.41	573.26	0.001
Outcome	1	5761.60	1566.52	0.001
Age $\times$ Ability	2	17.06	4.64	0.01
Age $\times$ Effort	2	60.22	16.37	0.001
Age $\times$ Outcome	2	40.29	10.96	0.001
Effort $\times$ Outcome	1	28.03	7.62	0.01
Age $\times$ Sex $\times$ Outcome	2	24.99	6.79	0.01
Age $\times$ SexS $\times$ SexO $\times$ Outcome	2	22.87	6.22	0.01

With one exception the results obtained for the main effects here are supportive of those obtained by Weiner and Peter. Outcome and effort are significant factors in both studies (a positive outcome and high effort receiving the greater reward). Also, neither study reports a significant effect for ability or for SexS (Weiner and Peter did not investigate SexO). The one exception concerns age where Weiner and Peter report a significant effect with the older children being more lenient in their judgments. The results obtained here are in the same direction (means of 0.81, 0.86 and 0.97 for the 9-, 10- and 15-year-olds respectively), but the differences do not approach significance.

This difference in the results may well be due to the greater age range used by Weiner and Peter although a graphical representation of their results (Weiner and Peter, Fig. 8) suggests that most of the age-related change they found does in fact take place within the age limits included here.

The results obtained by Salili as far as main effects are concerned, are also broadly supportive of Weiner and Peter with the exception of the significant main effect for ability referred to in the introduction.

Each of the three description variables (ability, effort and outcome) interact separately with age. The age $\times$ ability interaction shows that only the 15-year-olds give significantly greater rewards for lack of ability than they do for its presence ($\bar{X}$: high ability = 0.73, low = 1.21, 15 years only, standard error of difference (SEd) = 0.15). (Here, as throughout the report, all significant differences between cell means are significant at $P < 0.01$.) The two younger age groups did not give significantly different ratings to the two levels of ability. The same effect is reported by Weiner and Peter for their 16-18-year group while Salili reports no such effect. While these results do not contradict Weiner and Peter's interpretation—that only the more

cognitively mature subjects can operate the compensation schema, i.e., infer greater effort for success given little ability—evidence from elsewhere (Karabenick and Heller, 1976) shows that children as young as 6 years can operate the compensation schema. A more probable explanation would be that while younger children can use this schema it is not until 15 or so years of age that they actually do so in the evaluation of others. It is not yet clear why this should be.

Age also interacts with effort. Positive effort is increasingly rewarded with age with a significant increase taking place between 9 and 12 years ($\bar{X}$ s: 9 years = 1.5; 12 = 2.08 and 15 = 2.2; $SEd = 0.15$). Similarly lack of effort is punished more by the older groups with the significant change again occurring between the ages of 9 and 12 ($\bar{X}$: 9 = 0.11; 12 = 0.36 and 15 = 0.26, $SEd = 0.15$). Resultant effort scores (the absolute difference between the mean scores for positive and negative effort) show that effort is a more discriminating factor in the judgments of the two older groups of children (residual effort: 9 = 1.39; 12 = 2.44; 15 = 2.46). This result bears some similarity to that obtained by Weiner and Peter. They found a significant age $\times$ effort interaction that showed effort to be an increasingly significant factor in the judgments of children up to the age of 10-12 years but to become increasingly less so from then on. Such a reversal does not take place in the present study between the ages of 12 and 15 years. Salili also failed to find the reversal effect, effort again becoming increasingly important with age. After the age of 10 years the Iranian children gave increasingly greater rewards for positive effort and the same amount of punishment for the lack of it.

In the present study age also interacts with outcome while this is the case in neither Weiner and Peter nor in Salili. With negative outcomes the judgments become increasingly less severe with age with the significant difference being between the 9-year-olds and the others ($\bar{X}$ s: 9 = -1.22; 12 = -0.74; 15 = -0.61, $SEd = 0.17$). For positive outcomes the 9-year-olds again give the most extreme rating but those obtained from the 15-year-olds are now slightly more extreme than those from the 12-year-olds ($\bar{X}$ s: 9 = 2.83; 12 = 2.45; 15 = 2.55; $SEd = 0.17$). The only significant difference here lies between the 9- and 12-year olds. The resultant outcome scores reflect the way in which outcome becomes less important with increasing age but as can be seen from the final section of Table 2 outcome is, overall for each age group, the most discriminating of the three description variables of ability, effort and outcome.

As well as interacting with age, effort and outcome also interact with each other. Both Weiner and Peter and Salili report significant interactions here. The present results show that children differentiate between positive and negative effort to a greater extent when the outcome is negative than when it is positive. Similar results are obtained by Salili but with one important difference. In the present study positive outcome with negative effort is rewarded more than negative outcome with positive effort (1.68 and 0.32 respectively) while the opposite applies to Salili's data (0.5 and 1.75 respectively, figures extrapolated from Salili, Fig. 5). This would reflect the apparently greater emphasis placed on outcome in English *versus* Iranian culture.

Weiner and Peter do not discuss their effort $\times$ outcome interaction but instead a significant age $\times$ effort $\times$ outcome interaction. Salili also finds this interaction but it is not obtained in the present study. Weiner and Peter's interaction provides further evidence for their regression effect with effort being maximally important for the 10-12-year-olds. Salili states that they have no evidence to support the regression effect and the present results, so far, also provide no support.

The three remaining interactions all include the variable SexS. The SexS $\times$ effort interaction shows, first of all, that girls use effort information more than boys in making their evaluations (residual effort scores for girls and boys respectively = 2.31 and 1.89). Second, this difference is primarily due to their different evaluations of

positive effort ($\bar{X}$ s: girls/E+ = 2.07; boys/E+ = 1.79; girls/E- = -0.24; boys/E- = -0.01; SED = 0.14 for differences between sexes). Neither Weiner and Peter nor Salili report any such effect.

The Age $\times$ SexS $\times$ Outcome and Age $\times$ SexS $\times$ SexO $\times$ Outcome interactions clearly complement one another. Figure 1 shows the residual outcome scores derived from both interactions together with the cell means from the age $\times$ SexS $\times$ SexO $\times$ outcome interaction. The residual scores for the age $\times$ SexS $\times$ outcome interaction show clear regression of the type found by Weiner and Peter (outcome minimally important for the 10-year-olds) but for the girls only. For the boys outcome information has a decreasing effect on the judgments made throughout the age range studied. The residual scores obtained from the higher order interaction show that the girls produce essentially the same pattern when they are judging either girls or boys. For boys, however, the decline in importance of outcome information is more marked when they are judging girls than when judging boys.

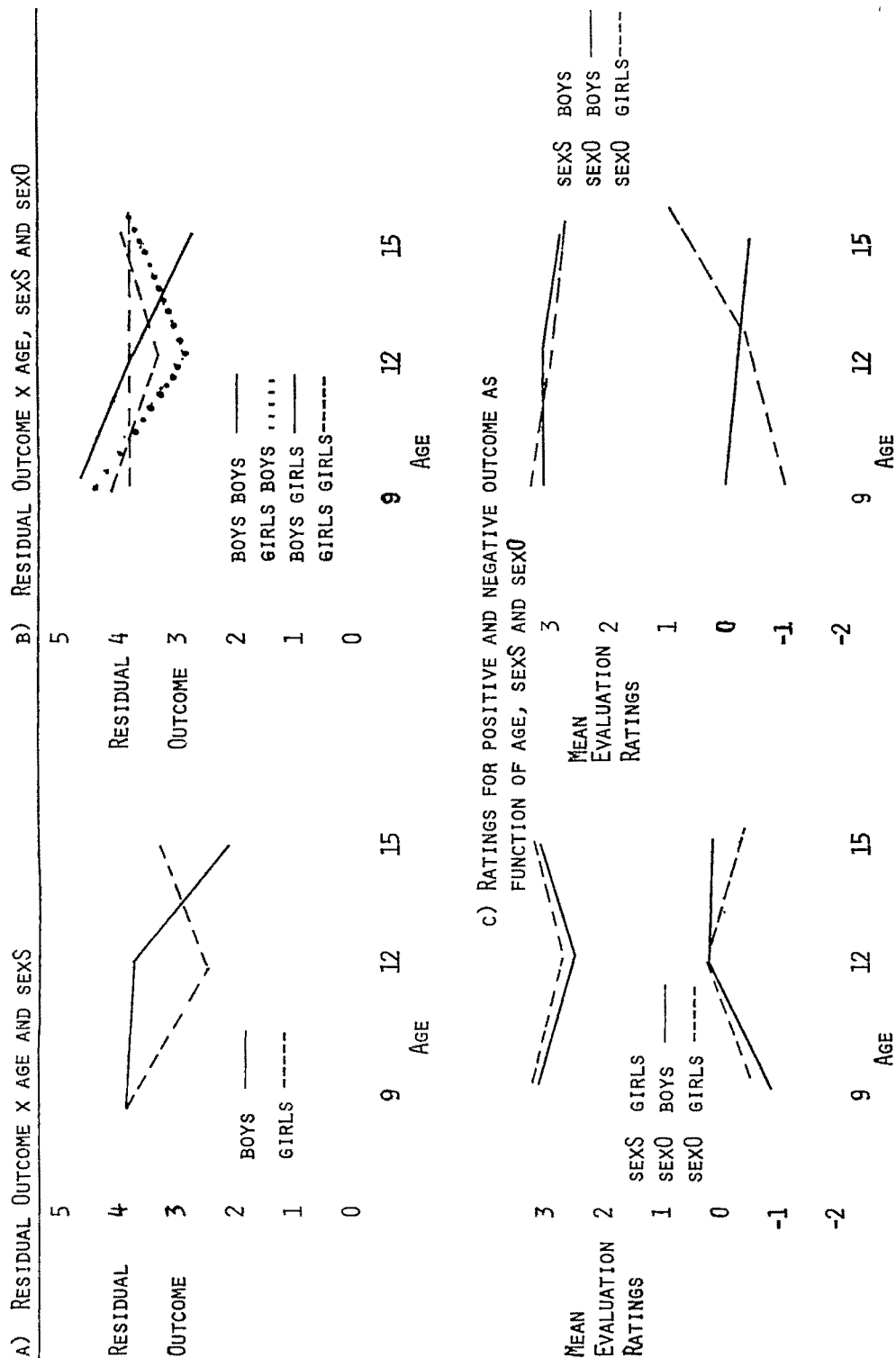
TABLE 2
RESIDUAL SCORES FOR ABILITY, EFFORT AND OUTCOME AS FUNCTION: AGE, SEXS AND SEXO

SexS	SexO	Age	Ability	Effort	Outcome
Girl	Girl	9	-0.18	1.43	3.93
		12	0.16	3.24	2.84
		15	-0.75	2.38	3.63
Girl	Boy	9	0.45	1.30	4.25
		12	0.04	2.56	2.59
		15	-0.69	2.91	3.29
Boy	Girl	9	0.05	1.15	4.58
		12	0.14	2.01	3.66
		15	-0.24	2.03	2.26
Boy	Boy	9	0.25	1.68	3.43
		12	-0.40	1.95	3.68
		15	-0.26	2.51	3.46
Overall		9	0.16	1.39	4.05
		12	-0.07	2.44	3.19
		15	-0.48	2.45	3.16

The bottom half of Figure 1, showing the cell means from the higher order interaction, clarifies this sex difference. Girls' judgments of positive and negative outcomes are near mirror images of each other. The oldest and the youngest girls give the most extreme scores. Boys' judgments of positive outcomes do not differentiate between the sexes and there is only a very marginal age-related change. For negative outcomes, however, boys' judgments of boys and girls differ markedly from each other. Nine-year-old boys punish girls significantly more for failure than they do boys ($\bar{X}$ s: boys/girls = -1.66; boys/boys = -0.73; SED = 0.33). Fifteen-year-old boys, however, punish their own sex significantly more severely, in fact they actually reward girls for failure ($\bar{X}$ s: boys/girls = 0.10; boys/boys = -1.10; SED = 0.33). Neither Weiner and Peter nor Salili have examined the effect of the person being judged or discussed significantly effects involving the sex of the judge.

Table 2 shows the resultant scores for ability, effort and outcome as a function of subject age and sex and the sex of the person being judged. There is only one case where the residual effort score is greater than that for residual outcome—12-year-old girls making judgments about other girls. The same subjects produce essentially equal resultant effort and outcome scores when judging boys. Otherwise outcome is always the most discriminating factor and ability always the least so.

FIGURE ONE:



DISCUSSION

Generally, as would have been expected, the results of the present study are in closer agreement with those obtained by Weiner and Peter than by Salili *et al.* In neither the US nor England is ability a significant factor in the judgments of children while it is in Iran. The oldest children in both western studies actually rewarded lack of ability. For the factor of effort the present results are not clearly supportive of either previous study. Effort becomes increasingly important with age and there is no sign of the regression effect obtained by Weiner and Peter. While effort is the single most important determinant of the judgments of the Iranian children this is only the case for the 10-12 year-olds in the US and for 12-year-old girls judging other girls in England.

The only support found for the Weiner and Peter regression effect is limited—outcome is minimally important for girls at 12 years but more important for both older and younger girls. Salili does not provide supportive data either. It is difficult, given the present data, to claim these and other differences between the three studies as cross-cultural. There are many differences of detail between the studies and it is quite possible that replications of the present study in other English schools would produce yet further differences in the results that may be more marked than those presently obtaining between England, the USA and Iran. In particular, the present study only related to children's evaluations of the performances of others on a test of mathematics. Various school subjects are seen differently in terms of their desirability and difficulty and these differences are sex-linked with mathematics being seen as being particularly difficult by girls (Duckworth and Entwistle, 1974; Maccoby and Jacklin, 1974). This last point needs to be borne in mind during the following discussion of the differences between the sexes observed here.

These results best serve to illustrate the complexity of developing differences between the sexes with respect to achievement motivation. While earlier studies (conducted in the USA) have shown girls to neglect effort as a cause of failure (Dweck and Reppucci, 1973; Nicholls, 1975) the present results shown that girls place a higher value than do boys on positive effort and, as indicated by the residual effort scores, use the information supplied concerning effort to a greater extent than the boys. These two sets of results are not necessarily contradictory. It is quite possible for girls to express the view that the exertion of a high level of effort is something to be valued and rewarded while at the same time not believing that effort, or the lack of it, is particularly responsible for their own successes and failures. It must also be remembered that girls, unlike boys, attach increasing importance to the actual outcome of another's achievement strivings between the ages of 12 and 15. Girls are certainly not characterised by the present results as having an over-riding concern with personal effort.

The development of an understanding of female attitudes towards achievement cannot be advanced without full consideration of male attitudes towards female achievement (and vice versa). It is in this respect that the present study supplies its most interesting results. While girls demonstrate a Weiner and Peter style regression effect with their judgments of the actual outcomes obtained by both boys and girls, the boys' judgments revealed different developmental trends with respect to each sex. When boys judge other boys the value placed upon the outcome obtained remained equal across the age range studied. However, when boys judge girls, the outcome obtained apparently becomes less valued with age. It will be recalled that this effect is primarily due to the judgments made of negative outcomes, with 9-year-old boys being most severe on girls who failed while 15-year-old boys were most charitable. These results for the older boys are supportive of earlier studies (e.g., Deaux and Taynor, 1973) showing that while men are rewarded for success women tend to be

excused for failure. The results of the present study show that such attitudes are acquired between the ages of 9 and 15 years.

The results of the present study have important implications for the future development of an attributional theory of achievement motivation, particularly concerning the development of sex differences. The majority of studies to date that have investigated aspects of an attributional theory of achievement motivation have essentially ignored the context in which attributions for success and failure are made. There have been some exceptions to this. Dweck (1979) reports studies showing that the typical patterns of negative and positive feedback obtained by male and female pupils in American schools differ and that these differences may relate to the typical attributions made by boys and girls for success and particularly for failure. Dweck and Bush (1976) have further shown that while girls may generally make maladaptive responses to failure feedback delivered by adults this is not the case when identical feedback is delivered by a peer. Indeed, when the feedback is delivered by peers, particularly male ones, it is the boys rather than the girls who make the maladaptive responses. The present results suggest that the girls' responses in this latter case may well be due to the condescending attitude of the older males towards female failure (i.e., girls have possibly learned not to always take feedback from boys at face value) but the Dweck and Bush study would need replication in the same context as the present study before this suggestion can be given serious consideration.

By demonstrating the way in which the evaluative judgments investigated here vary as a function of the overall context (i.e., broad comparisons between England, the USA and Iran) and the sex of both judge and judged, the importance of investigating judgments of achievement and the development of these within specified contexts is underlined. In any classroom it is the teacher and classmates of a particular pupil who will largely determine the context of that particular pupil's judgments. If the development of those judgments is to be understood, then future research must concentrate on elaborating the nature of the relationships that exist between those developing judgments and those of other relevant people.

REFERENCES

- ATKINSON, J. W., and RAYNOR, J. O. (1974). *Motivation and Achievement*. Washington, D.C.: V. H. Winston and Sons.
- CRANDALL, V. C. (1969). Sex differences in expectancy of intellectual and academic reinforcement. In SMITH, C. P. (Ed.), *Achievement-Related Motives in Children*. New York: Russell Sage Foundation.
- DEAUX, K., and TAYNOR, J. (1973). Evaluations of male and female ability: bias works two ways. *Psychol. Rep.*, 32, 261-262.
- DUCKWORTH, D., and ENTWISTLE, N. J. (1974). Attitudes to school subjects: a repertory grid technique. *Br. J. educ. Psychol.*, 44, 76-78.
- DWECK, C. S. (1979). Achievement. In LAMB, M. E. (Ed.), *Social and Personality Development*, London: Holt, Rinehart and Winston.
- DWECK, C. S., and BUSH, E. S. (1976). Sex differences in learned helplessness (I). Differential debilitation with peer and adult evaluators. *Dev. Psychol.*, 12, 147-156.
- DWECK, C. S., and GILLIARD, D. (1975). Expectancy statements as determinants of reactions to failure: sex differences in persistence and expectancy change. *J. Pers. soc. Psychol.*, 32, 1077-1084.
- DWECK, C. S., and REPPUCCI, N. D. (1973). Learned helplessness and reinforcement responsibility in children. *J. Pers. soc. Psychol.*, 25, 109-116.
- ICKES, W., and LAYDEN, M. A. (1978). Attributional styles. In HARVEY, J. H. ICKES, W., and KIDD, R. E. *New Directions in Attributional Research. Volume Two*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- KARABENICK, J., and HELLER, K. (1976). A developmental study of effort and ability attributions. *Dev. Psychol.*, 12, 559-561.
- KELLEY, H. H. (1972). Causal schemata and the attribution process. In JONES, E. E., KANOUSE, D. E., KELLEY, H. H., NISBETT, R. E., VALINS, S., and WEINER, B., *Attribution. Perceiving the Causes of Behaviour*. Morristown, New Jersey: General Learning Press.
- MACCOBY, E. E., and JACKLIN, C. N. (1974). *The Psychology of Sex Differences*. Stanford, California: Stanford University Press.

- NICHOLLS, J. G. (1975). Causal attributions and other achievement-related cognitions: effects of task outcomes, attainment value and sex. *J. Pers. soc. Psychol.*, 31, 379-389.
- SHULTZ, T. R., BUTKOWSKY, L., PEARCE, J. W., and SHANFIELD, H. (1975). Development of schemes for the attribution of multiple psychological causes. *Dev. Psychol.*, 11, 502-510.
- SALILI, F., MAEHR, M. C., and GILLMORE, G. (1976). Achievement and morality: a cross-cultural analysis of causal attribution and evaluation. *J. Pers. soc. Psychol.*, 33, 327-337.
- SMITH, M. C. (1975). Children's use of the multiple sufficient causal schema in social perception. *J. Pers. soc. Psychol.*, 32, 737-747.
- WEINER, B. (1974). *Achievement Motivation and Attribution Theory*. Morristown, New Jersey: General Learning Press.
- WEINER, B., and KUKLA, A. (1970). An attributional analysis of achievement motivation. *J. Pers. soc. Psychol.*, 15, 1-20.
- WEINER, B., and KUN, A. (1976). The development of causal attribution and the growth of achievement and social motivation. In FELDMAN, S., and BUSH, D., *Cognitive Development and Social Development*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- WEINER, B., and PETER, N. (1973). A cognitive developmental analysis of achievement and moral judgements. *Dev. Psychol.*, 9, 290-309.

(Manuscript received 21st January, 1980)

THE HERITABILITY OF FLUID, g_f , AND CRYSTALLISED, g_c , INTELLIGENCE, ESTIMATED BY A LEAST SQUARES USE OF THE MAVA METHOD

By R. B. CATTELL

(Psychology Department, University of Hawaii at Manoa)

SUMMARY. The heritability of the two second order general intelligence factors g_c and g_f , was studied by the MAVA design, employing 470 brothers reared together, 94 identical twins reared together, 124 fraternal twins reared together, and 1,543 boys (fluid intelligence) and 2,579 (crystallised intelligence) unrelated. All were in the 12-18 age range.

The model and the empirical data on four constellations give seven simultaneous equations for predicting expected values of the concrete, empirical variances from seven abstract, unknown variances and correlations. The best fit was obtained, using a least squares solution by computer program, CAMALS, and evaluated by a Q statistic analogous to chi-square. Two models offering reduced, parsimonious modifications of the full MAVA, were also tried but MAVA gave the highest index of good fit.

Substantial differences exist between the heritability of g_f and g_c , alike in the present and in earlier MAVA studies, using differing methods of solution. H for crystallised intelligence is highest for *within* family variance and the mean value stands at only 0.45 for the general population—lower still by the present method. Fluid intelligence, on the other hand, has the highest heritability *between* families, reaching a mean of 0.92. Its general population value is 0.58 closely approaching the 0.60 which has tended to be the lower end of the range by the twin method. (The reason for the usual higher twin method values is exposed by the finding of higher sib than twin within family threptic variance.) Some support is found for a *negative* genothreptic correlation between families, fitting the socio-psychological law of environmental coercion toward the mean.

1. DIFFERENCES FROM PREVIOUS STUDIES

THE present research offers certain technical advances over previously published analyses of intelligence inheritance, though its results conflict appreciably with those of others and earlier results of our own (Cattell *et al.*, 1981). The technical advances relative to one or another of past studies are: (1) The measurement of intelligence, in accordance with the evidence over the past decade (Horn and Cattell, 1966; Cattell, 1971; Horn, 1972, 1976; Haktian and Cattell, 1974; Stankow and Horn, 1980; and Vandenberg *et al.*, 1980), that *two* broad factors span the cognitive judgment domain by separate measure of fluid (g_f) and crystallised (g_c) intelligence. Traditional tests like WAIS and WISC are uncertain mixtures of these. Fluid and crystallised intelligence differ in many properties, age curves, IQ sigma, physiological associations and separate fluctuation over time (Cattell, 1971; Horn, 1972, 1976; Eysenck and Fulker, 1979). (2) The use of the MAVA (Multiple Abstract Variance Analysis) method instead of the twin method. (3) The use of a least-squares solution for the simultaneous equations involved, in place of the OSES method previously used on these data. These differences are explained in what follows.

2. THE PSYCHOMETRIC INSTRUMENTS AND THE MAVA DESIGN

The measures used for g_f and g_c are well known and factor-validated (concept validated) and need little digression in description. For g_f we used Forms A and B of Scale 2 of the IPAT Culture Fair Intelligence Test (1959, 1973). Each form covers 12 series, 14 classifications, 12 matrices, and 8 topology items, thus, with both forms representing 92 items and an actual testing time (i.e., without instruction time) of 20 minutes. For g_c both the B factor scales—from the A and B equivalent forms of the High School Personality Questionnaire (HSPO, Cattell and Cattell, 1969)—were

used. These together cover 20 items which are variously classification, analogies, synonyms, problems, opposites, in the verbal and numerical media. Here the test is untimed. The reliabilities and validities of both the Culture Fair and the HSPO are given in the respective handbooks, and elsewhere (Cattell, 1979).

These tests were administered to boys between the ages of 12 and 18 under the usual school conditions. The total number of cases was distributed among family constellations as shown in Table 1.

TABLE 1
SIZES OF SAMPLES USED FOR MEASUREMENT

Symbol		Number of Families	Number of Individuals
ITT	Identical twins raised together	47	94
FTT	Fraternal twins raised together	61	124
ST	Sibs (brothers) raised together	233	470
GP(2)	General population (unrelated raised apart), national standardisation. (IPAT, 1973)		For crystallised intelligence 2579
			For fluid intelligence 1543
*GP(1)	General population (unrelated raised apart), local		394

* The 'local' sample consists of one brother from each of the ST group plus enough from the same classrooms to bring up to this sample size. The present least squares solution did not use it, but others did. Note that the number of individuals is not exactly twice the number of families, because of the few cases with more than two in the 12-18 year range.

Turning from the psychometric basis to the experimental design we encounter a more complex matter. The Multiple Abstract Variance Analysis (MAVA) design (Cattell, 1960, 1963) was adopted because of its various advantages over the twin method, discussed elsewhere (Cattell, 1960; Cattell, Klein and Schuerger, 1980), but briefly to be mentioned here. Its generic relation to the methods of biometrical analysis as used in controlled animal breeding, etc., have been discussed by Jinks and Fulker (1970). Briefly its advantages over the twin method are that (1), it does not assume the same threptic variance for sibs as for twins, much evidence indicating that such assumption of quantitative identity of environmental divergence within sib and twin pairs is quite erroneous, and (2), it yields *between* family *and within* family genetic and threptic variances, as well as correlations (covariances) of genetic and threptic effects. In the present case not all of these advantages could be gained because of a limited supply of equations from a limited number of family constellations. In particular we would have liked to enter with different terms—unknowns—for the within family environmental (threptic) variance of identical and fraternal twins, but with our resources, stopping at six concrete, observed variances, we had to be content for the present with the advance of distinguishing, more importantly, the twin and the sib within family environmental variances.

It will be noted that the term *threptic* is used here and in a dozen other publications by various writers in the behaviour genetic field in place of *environmental*. The term *environmental* is ambiguous and the source of constant confusion in discussion for it may mean (as in some behaviour genetics) (a) the part of a trait that is *environmentally picked up* and now belongs to the existing personality, or (b) the *external* environmental variance which is the *cause* of the threptic effect. (*Threptic*, from the Greek for 'nourished' (Aristotle) is thus a neater substitute for the circumlocution *the inner structurally produced effect from an external environmental influence*.) The distinction of threptic and environmental variances is particularly important

when we come to investigation by the *comparative MAVA* design of the learning mechanisms connecting threptic variances with physico-social environmental variances (income, climate, years of education, etc.).

The essence of the MAVA method is that it sets up equations, according to a generally accepted model, for the concrete experimental variances to be expected for a number of family constellations, e.g., sibs reared together, identical twins reared together, sibs reared apart, fraternal twins reared together, unrelated children reared together, etc. It causally explains these observed *concrete variances* by breaking them down in terms, in the equations, representing the contributions of specified *abstract variances*. This is shown in Table 2 with such terms as σ_{wg}^2 for within family genetic variance, $\sigma_{wt.s}^2$ for within family threptic variance between sibs, σ_{bg}^2 for between family genetic variance, and so on. On the given data it then aims to solve as many simultaneous equations as are presented by the number of constellations, reaching values for the same number of abstract unknowns (σ_{wg}^2 , σ_{bt}^2 , etc.) as there are observed variances.

The problems that have presented themselves in research by this method have been, (1) to obtain enough family constellations to solve for the number of unknowns one would ideally like to know, in an adequate model of genetic and environmental interaction, (2) to obtain samples for the concrete variances that are of sufficient size to present the summation of sampling error blurring the results, and (3) to solve the simultaneous equations by the most efficient method. These are next considered.

3. SAMPLE SIZES AND OTHER ISSUES IN THE MAVA MODEL

Table 1 shows the samples that were measured on the g_f and g_e batteries above, and the abbreviations that will be used as a key henceforth. All were males, for identical twins have to be same sex and the toss gave males, so we are comparing brothers with same sex (male) identical and fraternal twins. In accordance with the canons of science of meeting one unknown at a time it seemed far better to leave the analysis for female twins to another investigator rather than mix sexes initially. The distinction of fraternal from identicals was made on the usual bases, including measures on standard colour photos against a frame, finger prints, and especially the information the parents obtained from the doctor at the child's birth. A more detailed account of the twin selection is preserved by Klein (University of Hawaii). Nevertheless, as typical examination of past studies has shown, if there is error it is likely to be in the direction of including one or two fraternal as identicals.

The model used by MAVA (in the present case where we have only four constellations to work with (Table 1)) is shown in Table 2. Note that we can use both within and between variances for each group, since they are statistically independent and, although the general population is composed largely of 'natural families' (sibs raised together) we can also use σ_{GP}^2 , which would ideally be twice the sum of σ_{ST}^2 and σ_{BNF}^2 but is here experimentally independent.

We had two different sample estimates of the general population (unrelated raised apart) variance: (1) $\sigma_{GP(1)}^2$ consisting of boys taken at random for each pair of brothers (but not from twins of either kind, for the reason just given), together with 161 boys from the same classrooms as the brothers, giving 394 in all, and (2) a less local, but far larger sample from the standardisation of the HSPQ (Cattell and Cattell, 1969) and the Culture Fair Scale 2 (IPAT, 1959, 1973) (which, taking boys alone, give $\sigma_{GP(2)}^2$ totals of 1,543 and 2,579 respectively). We have used only the larger sample here, but readers are given both to try if they wish.

The equations in Table 2 are taken from those published earlier (Cattell, 1960) where their rationale is explained. As will be seen the assumptions are an additive

TABLE 2

THE MAVA EQUATIONS, FROM CONCRETE TO ABSTRACT VALUES, FOR SEVEN RELATIONSHIPS AND SEVEN UNKNOWNNS

(1) σ_{ITT}^2	$= \sigma_{wt.t}^2$
(2) σ_{FTT}^2	$= \sigma_{wg}^2 + \sigma_{wt.t}^2 + 2r_{wgt} \sigma_{wg} \sigma_{wt.t}$
(3) σ_{ST}^2	$= \sigma_{wg}^2 + \sigma_{wt.s}^2 + 2r_{wgs} \sigma_{wg} \sigma_{wt.s}$
(4) σ_{BNF}^2	$= \sigma_{wg}^2 + \sigma_{wt.s}^2 + 2\sigma_{bg}^2 + 2\sigma_{bt}^2 + 2r_{wgt} \sigma_{wg} \sigma_{wt.s} + 4r_{bgbt} \sigma_{bg} \sigma_{bt}$
(5) σ_{ITTF}^2	$= 2\sigma_{wg}^2 + \sigma_{wt.t}^2 + 2\sigma_{bg}^2 + \sigma_{bt}^2 + 2\sqrt{2}r_{wgt} \sigma_{wg} \sigma_{wt.t} + 4r_{bgbt} \sigma_{bg} \sigma_{bt}$
(6) σ_{FTTF}^2	$= \sigma_{wg}^2 + \sigma_{wt.t}^2 + 2\sigma_{bg}^2 + 2\sigma_{bt}^2 + 2r_{wgt} \sigma_{wg} \sigma_{wt.t} + 4r_{bgbt} \sigma_{bg} \sigma_{bt}$
7) σ_{GP}^2 (or σ_{UA}^2)	$= \sigma_{wg}^2 + \sigma_{wt.s}^2 + \sigma_{bg}^2 + \sigma_{bt}^2 + 2r_{wgs} \sigma_{wg} \sigma_{wt.s} + 2r_{bgbt} \sigma_{bg} \sigma_{bt}$

KEY—*Abstract variances*: *wt.t* = within family threptic, twins; *wg* = within family genetic, *wt.s* = within family threptic, sibs; *bg* = between family genetic; *bt* = between family threptic; *r_{wgt}* = genothreptic correlation of within family genetic and threptic deviations; *r_{bgbt}* = genothreptic correlation, between family mean deviations threptic and genetic; *ITT* = identical twins raised together (within); *FTT* = fraternal twins raised together (within); *ST* = sibs raised together; *BNF* = between natural families, i.e. means of sibs raised together; *ITTF* = between identical twins raised together families; *FTTF* = between fraternal twins raised together families; *GP* = general population (or *UA*, unrelated individuals raised apart).

action of genes and environment with recognition that there may be covariance (quantified in *r_{wgt}* and *r_{bgbt}*) between them for any of the various reasons analysed earlier (Cattell, 1960, 1963).

In anticipating that these can be directly solved as seven simultaneous equations for seven unknowns one would overlook (a) that they are not simple linear equations, and (b) that there could turn out to be hidden dependencies, and in this case it prevents $\sigma_{bg}^2 + \sigma_{bt}^2 + 2r_{bgbt} \sigma_{bg} \sigma_{bt}$ from being broken apart. In order to solve, therefore, we brought in an assumption that can be firmly made on known genetic principles, namely, that the unknown σ_{bg}^2 will bear a fixed relationship (as calculated below) to the solvable σ_{wg}^2 .

4. BASES OF THE STATISTICAL ANALYSIS: THE OSES AND THE LEAST SQUARES APPROACHES

In an earlier analysis of the data, which also used this assumption (Cattell *et al.*, 1981) we attempted an alternative path to solution from the present. This has been dubbed the OSES solution for *Overlapping Simultaneous Equation Sets* solution. When no unique solution is found possible for all equations together, through dependencies, alternative quadratic solutions (affecting *r_{bgbt}* here), and having one more equation than there are unknowns, it may happen, as it does here, that solutions can be found through smaller sets (in this case 5 and 6 sets, of which theoretically there are respectively 21 and 7). In the OSES presentation we averaged, for each unknown, the solutions found across as many sets as yielded solutions for it. The uneven weighting and lack of elegance in this causes the OSES method to appeal to statisticians less than a maximum likelihood or least squares method, but we have published both solutions because we feel the OSES, despite this weakness, has some value in permitting one to see by what weightings of sources different results are obtained.

In any case a digression must be made in the present least squares, as in the OSES, to explain the derivation of σ_{bg}^2 from σ_{wg}^2 . It is a firm conclusion from Mendelian laws that without dominance effects and with random mating in the population σ_{bg}^2 literally equals σ_{wg}^2 . Naturally, in human society random mating cannot be assumed, and we must proceed to the more precise relation where

$$\sigma_{bg}^2 = \sigma_{wg}^2 \left(\frac{1+r_{ga}}{1-r_{ga}} \right) \quad (8)$$

where r_{ga} is the assortive mating correlation for the *genetic* fraction of husband and wife scores.

There are two problems in inserting a value for r_{ga} , namely, in getting the phenotypic correlation from which r_{ga} is derived, and in deriving it. As to the first, Jensen (1978) records assortive mating correlations for intelligence ranging from +0.15 to +0.76 and others from 0.25 to 0.75. Vandenberg's (1980) survey is very similar. Cattell and Willson (1938) in England found a figure in the upper half of this range. Cattell and Nesselroade (1967) found in the mid-west a value of 0.31. In so far as one may use indirect evidence Cattell and Nesselroade found an average husband-wife r of +0.20 over 17 different personality factors and Woliver (1979) also one of +0.20 over 10, but it could be argued that intelligence being more relevant to social encounter and congeniality, its value would be higher (at the phenotypic level). Thus one might take a compromise of these results and Jensen's in the region of +0.30 to +0.40. However, our estimate of r_{ga} has taken account also of the likelihood (1) that assortive matings seem to have declined over the last decade, (2) that it is higher in a more class stratified society (as in the Cattell and Willson data from a settled English community) than would be likely in our ethnically mixed Chicago and Cleveland data, and that (as Jensen mentions) there are suggestions that it is lower in lower social classes, (3) that, as Rao and Morton (1970) have argued, the genetic correlation is likely to be lower than the phenotypic correlation. Concerning this last, however, the present writer (1981) has given some demonstration that except with quite unusual assumptions the most probable value for the genetic assortive r_{ga} is actually the same as for the phenotypic r_{pa} . Nevertheless, in deference to the above, the calculations were tried out with a range of assumed values somewhat lower, namely at 0 +0.10 and +0.25. These have turned out well in covering the required range, and our position is supported by Eaves' (1973) finding that assortive mating is essential to good fit in the case of intelligence inheritance data.

5. THE OBTAINED NUMERICAL SOLUTIONS FOR THE ABSTRACT VARIANCES AND CORRELATIONS

The design for analysis here was to take all seven equations, but to substitute for σ_{bg}^2 the value $k\sigma_{wg}^2$, where k is one of the three ratios developed from putting $r_{ga} = 0$, 0.10 and 0.25 in equation 1. This means obtaining a least squares fit of six unknowns to the seven equations. A computer program for this purpose indexed as CAMALS (Cattell's MAVA by Least Squares) was constructed by Professor D. C. Rao. The subsequent tables are from Dr. Rao's program and design for evaluation, but he is in no way responsible for the present writer's use of the program and the conclusions here drawn. However, it was Dr. Rao's suggestion that we try out, in addition to the full MAVA model, two more parsimonious models, one in which *no* genothrepic correlations (r_{wgwt} and r_{bgbt}) are assumed (leaving only σ_{wg}^2 , σ_{wt}^2 , σ_{wts}^2 and σ_{bt}^2) and one in which *no* genetic contribution is considered at all!

The fit of the model is examined by how well the true (observed) concrete variances are restored from the abstract variances reached from the collective MAVA equations. Table 3 takes the opportunity in presenting the concrete variances obtained in the experiments to present also the above expected values from the least squares fit. The value Q at the bottom is an (inverse) measure of the goodness of fit and is obtained by dividing the residual sum of squares (the function of the difference of concrete and expected values) by the square of the expected general population variance. This has a general resemblance to chi-square but cannot be directly so

TABLE 3
GOODNESS-OF-FIT OF THE MAVA MODEL TO FLUID (g_f) AND
CRYSTALLISED INTELLIGENCE (g_c)

Variance	Sample Size n_i	g_f		g_c	
		Obs.	Exp.	Obs.	Exp.
σ_{ITT}^2	94	58.99	58.31	2.47	2.57
σ_{FFT}^2	124	82.03	91.29	5.21	6.69
σ_{IT}^2	470	147.33	143.42	5.66	5.09
σ_{BITTF}^2	44	341.29	339.65	17.60	17.75
σ_{BFTTF}^2	124	288.12	280.62	17.12	15.48
σ_{BNF}^2	470	331.88	332.74	13.63	13.88
σ_{OP}^2	2579	236.96	238.08	9.42	9.48
Q*		0.48		8.95	
(Validity)					

evaluated. We have here taken any value less than 10 as a pretty satisfactory fit. Among nine personality factor source traits similarly studied elsewhere and the two intelligence factors this cut at 10 operates very conservatively, leaving only 69 per cent of the hypotheses (assortive mating correlations with concrete variances) accepted. As Table 3 shows, the MAVA model *as it stands* gives a good fit for both fluid and crystallised intelligence, but especially good for fluid.

A more radical perspective on the model would be to ask if the full MAVA is indeed the best model among those in the area by comparing its fit with those of (a) a model which fits the environmentalist's dream of *no genetic action at all* in intelligence and (b) a more parsimonious model than the full MAVA (but still not as parsimonious as (a)) in which *both* correlations of genetic and threptic deviations (r_{wgwt} and r_{bgbt}) and assortive mating correlation (r_{ga}) are dropped, as zero in value.

By our standard for Q the alternative models for g_f , on our present sample basis, are acceptable, but the full MAVA gives the best fit. In the case of crystallised intelligence, although the parsimonious and purely threptic explanations fail, the general MAVA itself is not as good as in the case of fluid intelligence.

In the case of crystallised intelligence the mean Q values for different inserted assortive mating (r_{ga}) values were 9.85, 9.78 and 9.48 respectively for $r_{ga} = 0$,

TABLE 4
TEST OF ALTERNATING MODELS FOR DETERMINATION OF FLUID INTELLIGENCE (g_f)

Estimates $\pm$ Measures of Error (ME) under various hypotheses						
Parameter	General MAVA		Parsimonious		Purely Threptic	
	Estimate	ME	Estimate	ME	Estimate	ME
$\sigma_{wrt. i}^2$	58.31	± 0.16	50.42	± 0.06	90.67	± 0.04
σ_{wg}^2	77.45	± 0.03	46.62	± 0.06	0	0
r_{wgwt}	-0.33	± 0.62	0	0	0	0
(Assortive mating, $\sigma_{wrt. s}^2$)	133.03	± 0.11	96.83	± 0.03	137.08	± 0.02
σ_{bg}^2	94.66	$\pm$	46.62	± 0.06	0	0
r_{ga} , at σ_{gt}^2	0	0	47.90	± 0.06	100.89	± 0.02
+0.10) r_{bgbt}	0	0	0	0	0	0
Q*	0.48		0.90		7.13	

The mean error ME conveniently giving the 50/50 range $\times$ standard error.

TABLE 5

TEST OF ALTERNATING MODELS FOR DETERMINATION OF CRYSTALLISED INTELLIGENCE (g_c)

Estimates $\pm$ Measures of Error (ME) under various hypotheses							
Parameter	General MAVA		Parsimonious		Purely Threptic		
	Estimate	ME	Estimate	ME	Estimate	ME	
(Assortive mating, r_{ga} at +0.25)	$\sigma^2_{wt. t}$	2.57	± 0.12	3.11	± 0.02	5.85	± 0.02
	σ^2_{wg}	0.96	± 0.19	3.17	± 0.02	0	0
	r_{wgwt}	1.00	0	0	0	0	0
	$\sigma^2_{wt.}$	1.62	± 0.14	1.91	± 0.03	4.65	± 0.02
	σ^2_{bg}	1.61	0	3.17	± 0.02	0	0
	σ^2_{bt}	5.14	± 0.04	1.22	± 0.06	4.83	± 0.02
	r_{bgbt}	-0.41	0	0	0	0	0
	Q^*	8.95		10.42		28.56	

The mean error ME conveniently giving the 50/50 range by $0.67 \times$ standard error.

$r_{ga} = +0.10$ and $r_{ga} = +0.25$, so the last was chosen. For fluid intelligence the corresponding values were 3.80, 3.03 and 4.28, indicating a most likely (but not significantly so) genetic assortive mating value of +0.10.

6. CALCULATION OF HERITABILITIES

Let us note in passing that the present MAVA studies are perhaps the first to demonstrate definitely that there is a substantial difference of twin and sib (brother) threptic variances. This finding does not appear convincingly in the present results where the mean of the two ratios shows them to be the same, due to the unusually reversed crystallised intelligence ratio. But in 21 diverse personality factor source traits, measured by both questionnaire and objective (OA) test batteries (Cattell, Klein and Schuerger, 1980; Cattell, Schuerger and Klein, 1980; Cattell, Schuerger, Klein and Kameoka, 1980) the $\sigma_{wt}^2 / \sigma_{wt}^2$ ratio centres on 1.6. Since practically all previous estimates of within family heritability (often simply called 'heritability') have been estimates of heritability, previously based on the twin method, existing surveys are significantly biased toward an overestimate of H , as a typical family or general population value.

It should be noted that there are two different bases for calculating H , whether within family, between family or general population heritability, which we have here called *gross* and *net*. The former (H_g for gross) divides by the total variance as it actually is; thus in case of population heritability:

$$H_{g.p} = \frac{\sigma_{wg}^2 + \sigma_{bg}^2}{\sigma_{wg}^2 + \sigma_{wt}^2 + 2r_{wgwt}\sigma_w\sigma_{wt} + \sigma_{bg}^2 + \sigma_{bt}^2 + 2r_{bgbt}\sigma_{bg}\sigma_{bt}} \quad (9)$$

The net ignores whatever covariance exists, on the ground that this might be considered a relative accidental feature, and thus calculates

$$H_{n.p} = \frac{\sigma_{wg}^2 + \sigma_{bg}^2}{\sigma_{wg}^2 + \sigma_{wt}^2 + \sigma_{bg}^2 + \sigma_{bt}^2} \quad (10)$$

The argument for H_n is that it deals simply with genetic and environmental variances *in themselves*, which may be more relevant for a biologist, or a psychologist concerned with learning and *nurturability* (Cattell, 1973a). But from a social psychological point of view, e.g., in studying family interactions or social status effects, the r_{wgwt} and

r_{bgbt} are far from being justly treated as 'accidental' and the H_g makes better sense. Their relation is simply that H_g is smaller if covariance is positive, while H_n is smaller if it is negative.

TABLE 6
THE THREE MAIN HERITABILITIES AND GENOTHEPTIC CORRELATIONS

	Fluid Intelligence (g_f)		Crystallised Intelligence (g_c)	
	Gross	Net	Gross	Net
$H_{w(t)}$ within family sibs	*44	0.37	0.25	0.37
H_b between family	1.0	1.0	0.29	0.24
H_p over total population	0.63	0.56	0.27	0.28
r_{wgt}		-0.33		1.00
r_{bgbt}		0		-0.41
$H_{w(t)}$ within family, twins,	28*	0.57	0.53*	0.27

* It is symptomatic of the conceptual difficulties with the twin method that (a) we have had to refrain here from calculating the covariance as that of the identical twin theptic variance and the sib or fraternal twin genetic difference, and (b) for alignment with the usual fraternal and identical comparison we have used the fraternal values rather than the sib, as elsewhere. Thus this is not calculated from the estimates, but by

$$H_{w(t)} = \frac{\sigma_{fTT}^2 - \sigma_{fTT}^2}{\sigma_{fTT}^2} \quad (11)$$

using the direct, concrete variances.

The heritabilities on both bases are calculated from the General MAVA (Best Fit) in Tables 4 and 5, for each of three situations, in Table 6. These are: *within family*, H_w , *between family*, H_b , and across the *whole population*, H_p . Separately, at the bottom, for those who wish to compare with results of earlier work using the twin method, the values are given from assuming that the between twin theptic variance is the same as that for ordinary sibs, i.e., $\sigma_{wg}^2 = \sigma_{ST}^2 - \sigma_{TT}^2$.

7. SUMMARY AND DISCUSSION

It is neither practicable nor our intention to discuss these results in relation to extensive reviews of previous work, very largely confined to the twin method. Excellent surveys (in terms also of methodological evaluations) exist in Erlenmeyer-Kimling and Jarvik (1963), Jencks *et al.* (1972), Jensen (1973), Eysenck (1971), Eysenck and Fulker (1979), Jinks and Eaves (1974), Loehlin (1965, 1978), Thompson (1967) and Vernon (1969) among others. Comparisons would in any case be difficult because of systematic differences in magnitudes from the assumption defined above in the twin method.

Nevertheless, it is desirable to put the present results in some perspective of agreement or disagreement with other results, and this can be done mainly through three MAVA method studies, those by Cattell *et al.* (1955) and Cattell *et al.* (1957) using solution of a complete set of simultaneous equations and a recent study by Cattell *et al.* (1981) which operated on the same data as the present research but by an OSES (overlapping simultaneous equation sets) analysis, which is less statistically

approved than the present. As far as the mainly twin method results are concerned mainly on intelligence tests using unknown mixtures of g_f and g_e , the values of the H s have ranged, in Mittler's survey (1971), from about 0.55 to 0.59, in Erlenmeyer-Kimling (see summary Cattell, 1971, p. 265) from 0.67 to 0.95, in Claridge *et al.* (1973) have intraclass r s essentially for g_f which give H in the same range; Nichols in Vandenberg (1965) gives 0.75. Jensen (1973) and Eysenck (1971) indicate that despite Burt's errors his Burt and Howard (1956) value close to 0.80 is central to most. Thus although recent in press studies run somewhat lower the twin method, with measures probably more g_e than g_f , and in terms of within family values (H_w) might be said to range around 0.70-0.80 as a central tendency.

The Cattell *et al.* (1957) and Cattell *et al.* (1955) had the tentativeness of pioneer studies, but their testing had the same g_f and g_e factor identification as the present data (the Culture Fair and the verbal classification, analogy, etc.) and their total samples were among the largest of published studies (647 and 962, respectively) until quite recently. Since these form a family of studies (differing only in that Cattell *et al.* (1955) included an 'unrelated boys raised together' constellation which this does not and Cattell *et al.* (1957) had both unrelateds together and sibs apart which this does not) it is useful to put their results side by side in Table 7.

TABLE 7
COMPARISON OF HERITABILITIES FROM FOUR MAVS ANALYSES
(For comparison all are taken as net H s not gross H s)

Study	Fluid (g_f) Intelligence				Crystallised (g_e) Intelligence			
	CSK '57	CKGS '81	Present	Mean	CBB '55	CKGS '81	Present	Mean
$H_{w(e)}$ within family, sibs	0.51	0.39	0.37	0.42	0.54	0.74	0.37	0.55
H_b between family	0.92	0.85	1.00	0.92	0.38	0.61	0.24	0.41
H_p over total population	0.62	0.57	0.56	0.58	0.47	0.61	0.28	0.45
r_{egwt}	0*	-0.39	-0.33	-0.24	-0.0*	-0.52	1.0	0.13
r_{hgt}	0.25*	-0.41	0	-0.05	-10*	-0.81	-0.41	-0.44

KEY: CSK = Cattell, Stice and Kristy (1957)
CKGS = Cattell, Klein, Graham and Schuerger (1981)
CBB = Cattell, Blewett and Beloff (1955)

* The CSK and CBB studies solved for as many equations as unknowns, but had arbitrary correlations with which however the variances are consistent (mean of two quadratic values). The CKGS used the OSES approach, averaging results from overlapping subsets of equations. In CBB the σ_{wt}^2 is that of fraternal twins and in CSK the mean of sibs and twins. In all, sib scores are corrected to the same age.

From a methodological point of view we would put most weight on the present, least-squares-fit program. Actually the rank order of its H values is no different from the mean of the others, namely, for fluid intelligence, in diminishing order, H_b , H_p , H_w and in crystallised H_w , H_p , H_b ; so a simple mean may be considered. Nevertheless, in crystallised intelligence its values are uniformly stepped down appreciably from others, which cannot be ascribed to difference of social populations (since CKGS, 1981, uses the same). The only flat contradiction in the table is the finding of a positive correlation of 1.0 for within family genetic and threptic deviations, in crystallised intelligence which we must at present leave as an anomaly for future research on methodological peculiarities to examine.

The main conclusions, in an over-all perspective, are:

(1) That the full MAVA model fits better than other models we have tried, for both fluid and crystallised intelligence.

(2) That the total across population heritability is noticeably lower than commonly found, largely by twin methods. It is these latter results which have previously led psychologists to think in the 60-80 per cent range of heritability values.

(3) This total heritability is here nevertheless roughly 60 per cent for fluid intelligence, though decidedly lower for crystallised intelligence (45 per cent). This size and direction of difference is consistent with all that has previously been written about the empirical supports and general features of the theory of fluid and crystallised intelligence (Horn and Cattell, 1966; Cattell, 1971; Horn, 1972, 1976).

(4) The heritability of fluid intelligence is particularly high (0.92) in the between-family means situation. This is strong support for the culture-fairness of the culture fair intelligence test designs (Cattell, 1979) and suggests that family background and social status have markedly less effect on the validly measured g_f than on the g_c intelligence. The, at first, surprising indication of a higher heritability ratio for crystallised than fluid intelligence *within* families suggests the hypothesis that the cultural accretions that combine with g_f to produce g_c have very little variance within families, i.e., that the H ratio is higher because the *threptic variance* of g_c within families is far smaller than it is between families and *not* because the genetic variance is large. Since g_f and g_c are in different units this expectation cannot be directly checked, but Tables 4 and 5 show the within family threptic variance altogether lower, compared to between family threptic variance in crystallised intelligence. Indeed, by the best method, the estimate of the between family threptic variance on culture fair (g_f) measures is zero! At the same time an ancillary and weaker hypothesis might need to be explored that fluid intelligence actually differs more at birth than would be expected among sibs, and this might have something to do with gestation conditions, birth order, age of parents at child's birth, etc. A general theory of *duo-threpsis* has been proposed (Cattell *et al.*, 1981) to bring more research attention to splitting the known threptic effect into two causes: *cultural-educational* as most commonly thought of for the 'stimulating environment for intelligence' and *biological*, acting in prenatal nutrition, birth injuries, and disease during growth. The latter would effect fluid intelligence (as now understood) more than crystallised intelligence, and lead to the lower within family H_w for g_f just discussed. This kind of investigation calls for supplementing MAVA with path-coefficient, causal analysis as now being pursued by Birkett and Cattell, 1978, Eysenck and Fulker, 1979, Rao and Morton, 1974, and Rice *et al.* (1978).

(5) Finally we come to the genothreptic correlations where, unfortunately, the standard deviations here, and other signs, justify little weight on most of the single study values. At least between families the evidence pretty uniformly indicates *negative* correlations, probably trivial for fluid intelligence, where in any case the effect of environment between families is itself very small, but quite appreciable (-0.44) for crystallised intelligence. This is consistent with the *law of coercion to the bio-social mean* (Cattell, 1960, 1963, 1971) which states that cultural influences (mainly in schools, neighbourhoods, and the mass media) tend to act towards coercing initial deviations, in whichever directions, toward the mean. In this case the lock-step of unselective education presses the backward to advance more rapidly but leaves the bright to mark time, while the mass effect may reward the transition of the intellectual discriminatory capacities of more gifted and trained high g_c individuals toward the communication levels of the average. (The lesser sigma of crystallised than of fluid IQs (Cattell, 1971) is further evidence of this.)

In retrospect on the above results the question arises whether the random errors of measurement and the sampling errors, e.g., putting some fraternal twins in an identical twin group, affect H values randomly or with bias. It is hard to handle the complexities in a paragraph. The present writer's conclusion elsewhere has been that reliabilities and validities, typically falling short of unity, affect H s little, since all concrete variances in Table 2 are in the main affected in the same ratio. However, most trait measures suffer from function fluctuation (Cattell, 1973b, p. 355; Horn, 1972) quite different from error of measurement. Strictly we should average several measures, over occasions, to define the trait level, and our failure to do so will bias toward higher threptic variances and lower H s. The second important systematic error, especially in intelligence tests, is the existence of what Vernon (1969) has well defined as *test sophistication*, and to which culture fair, g_f measures are quite as vulnerable as traditional intelligence tests (mainly g_c). Improvements of experiment will result here also from repeated measures now with the same type of test, but meanwhile this source of variance is falsely added to environmental effects on traits. Consequently it is realistic to consider the heritabilities published as being lower limits, though one may guess that the above biases in the error would not raise them very much. Finally, when all sources of error, experimental and methodological analytical, have been removed, one would still not expect identity of first order results from different experiments, for the H values will vary with racio-cultural ranges (see possible inferences from comparative MAVA (Cattell, 1973(b)). The apparently steady tendency for our American samples to fall to lower H s than from those in Britain could be due to the greater cultural uniformity in the latter.

ACKNOWLEDGMENTS.—The writer wishes to thank Dr. Klein and Professor Schuerger for data from their earlier article, Drs. J. Brennan and F. Ahern for sorting and computer programming the data, Dr. R. K. Graham for financial support in computing, and, pre-eminently, Dr. D. C. Rao for his penetrating discussions of design of analysis and his aid to Dr. Brennan in constructing the least squares analysis computer program by which these results were calculated.

REFERENCES

- BIRKETT, H., and CATTELL, R. B. (1978). Diagnosis of the dynamic roots of a clinical symptom by P-technique. *Multivariate exp. clin. Res.*, 3, 173-194.
- BURT, C. L., and HOWARD, M. (1956). The multifactorial theory of inheritance and its application to intelligence. *Brit. J. stat. Psychol.*, 9, 95-131.
- CATTELL, R. B. (1960). The multiple abstract variance analysis equations and solutions; for nature. nurture research on continuous variables. *Psychol. Rev.*, 67, 353-372.
- CATTELL, R. B. (1963). The interaction of hereditary and environmental influences. *Br. J. stat. Psychol.*, 16, 191-210.
- CATTELL, R. B. (1971). *Abilities: Their Structure, Growth and Action*. Boston: Houghton Mifflin.
- CATTELL, R. B. (1973a). Unravelling maturational and learning developments by the comparative MAVA and structured learning approaches. In NESSELROADE, J. R., and REESE, J. (Eds.), *Lifespan Developmental Psychology*. New York: Academic Press.
- CATTELL, R. B. (1973b). *Personality and Mood by Questionnaire*. San Francisco: Jossey Bass.
- CATTELL, R. B. (1979). Are culture fair intelligence tests possible and necessary? *J. Res. Dev. in Educ.*, 12, 3-13.
- CATTELL, R. B. (1981). *The Inheritance of Personality and Ability: Research Methods and Findings*. New York: Academic Press.
- CATTELL, R. B., BLEWETT, D. B., and BELOFF, J. R. (1955). The inheritance of personality. A multiple variance analysis determination of approximate nature-nurture ratios for primary personality factors in Q-data. *Am. J. human. Genet.*, 7, 122-146.
- CATTELL, R. B., and CATTELL, M. (1969). *The High School Personality Questionnaire*. Champaign, Ill.: Institute for Personality and Ability Testing.
- CATTELL, R. B., KLEIN, T. W., and SCHUERGER, J. M. (1980). Heritabilities by the MAVA method and objective test measures of personality traits, U.I.23, Capacity to Mobilize, U.I.24, Anxiety, U.I.26, Narcistic Ego, and U.I.28, Asthenia. (In press.)
- CATTELL, R. B., KLEIN, T. W., GRAHAM, R. K., and SCHUERGER, J. M. (1981). The heritability of fluid and crystallized general intelligence by the MAVA method and OSES analysis. (In press.)

- CATTELL, R. B., and NESSELROADE, J. R. (1967). Likeness and completeness theories examined by 16 PF measures on stably and unstably married couples. *J. pers. soc. Psychol.*, 7, 351-361.
- CATTELL, R. B., SCHUERGER, J. M., and KLEIN, T. W. (1980). The heritability of dominance, E; surgency, F; premisia, I; and self-sufficiency, Q₂ by questionnaire measurement and the MAVA method. (In press.)
- CATTELL, R. B., SCHUERGER, J. M., KLEIN, T. W., and KAMEOKA, V. (1980). The heritability of the clinically relevant personality factors: guilt proneness, O; parmia, H, ergic tension, Q₄, by questionnaire and the MAVA method. (In press.)
- CATTELL, R. B., STICE, G. L., and KRISTY, N. F. (1957). A first approximation to nature-nurture ratios for eleven primary personality factors in objective tests. *J. abnorm. soc. Psychol.*, 54, 143-159.
- CATTELL, R. B., and WILLSON, J. L. (1938). Contributions concerning mental inheritance. 1. Of intelligence. *Br. J. educ. Psychol.*, 8, 129-149.
- CLARIDGE, G., CARTER, S., and HUME, W. I. (1973). *Personality Differences and Biological Variations: a Study of Twins*. Oxford: Pergamon Press.
- COHEN, J. (1957). The factorial structure of the WAIS between early childhood and old age. *J. consult. Psychol.*, 21, 283-290.
- DEFRIES, J. C., JOHNSON, R. C., KUSE, A. R., MCCLEARN, G. E., POLOVINA, J., VANDENBERG, S. G., and WILSON, J. R. (1979). Family resemblance for specific cognitive abilities. *Behav. Genet.*, 9, 23-43.
- EAVES, L. J. (1973). Assortative mating and intelligence: an analysis of pedigree data. *Heredity*, 30, 199-210.
- ERLENMEYER-KIMLING, M., and JARVIK, L. F. (1963). *Science*, 142, 1477-1479.
- EYSENCK, H. J. (1971). *The IQ Argument*. New York: Library Press.
- EYSENCK, H. J., and FULKER, D. W. (1979). *The Structure and Measurement of Intelligence*. New York: Springer.
- HAKSTIAN, A. R., and CATTELL, R. B. (1974). The checking of primary ability structure on a broader basis of performances. *Br. J. educ. Psychol.*, 44, 140-154.
- HORN, J. L. (1972). The structure of intellect: primary abilities. In DREGER, R. H. (Ed.), *Multivariate Personality Research*. Baton Rouge: Claitor.
- HORN, J. L. (1976). Human abilities: a review of research and theories in the early 1970s. *Annual Rev. Psychol.*, 27, 437-485.
- HORN, J. L., and CATTELL, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *J. educ. Psychol.*, 57, 253-270.
- INSTITUTE FOR PERSONALITY AND ABILITY TESTING (1959, 1973). *The Culture Fair Intelligence Tests. Scales 1, 2 and 3*. Champaign, Ill.
- JENCKS, C. et al. (1972). *Inequality. A Re-assessment of the Effect of Family and Schooling in America*. New York: Basic Books.
- JENSEN, A. R. (1973). *Educability and Group Differences*. London: Harper.
- JENSEN, A. R. (1977). The current status of the IQ controversy. *Address at Australian Universities*.
- JENSEN, A. R. (1978). Genetic and behavioral effects of non-random mating. In OSBORNE, R. T., NOBLE, C. E., and WEYL, N. (Eds.), *Human Variation*. New York: Academic Press.
- JINKS, J. L., and EAVES, L. J. (1974). IQ and inequality. *Nature*, 248, 287-289.
- JINKS, J. L., and FULKER, D. W. (1970). Comparison of the biometrical genetical, MAVA, and classical approaches to the analysis of human behavior. *Psychol. Bull.*, 73, 311-349.
- KLEIN, T. W., and CATTELL, R. B. (1980). The heritability of fluid and crystallized intelligence as determined by the twin method. (In preparation.)
- LOEHLIN, J. C. (1965). Some methodological problems in Cattell's multiple abstract variance analysis. *Psychol. Rev.*, 72, 156-171.
- LOEHLIN, J. C. (1978). Heredity-environmental analyses of Jencks' IQ correlations. *Behav. Genet.*, 8, 415-436.
- MITTLER, P. (1971). *The Study of Twins*. Baltimore: Penguin.
- NICHOLS, R. C. (1965). The national merit twin study. In VANDENBERG, S. G. (Ed.), *Methods and Goals in Human Behaviour Genetics*. New York: Academic Press.
- RAO, D. C., and MORTON, N. E. (1970). IQ as a paradigm in genetic epidemiology. In *Genetic Epidemiology*. New York: Academic Press.
- RAO, D. C., and MORTON, N. E. (1974). Path analysis of family resemblance in the presence of gene-environment interaction. *Am. J. human Genetics*, 26, 767-772.
- RICE, J., CLONINGER, C. R., and REICH, T. (1978). Multifactorial inheritance with cultural transmission and assortive mating. 1. Description and basic properties of the unitary models. *Am. J. human Genetics*, 30, 618-643.
- STANKOW, L., and HORN, J. L. (1980). Human abilities revealed through auditory tests. *J. educ. Psychol.*, 72, 21-44.
- THOMPSON, W. R. (1967). Some problems in the genetic study of personality and intelligence. In HIRSCH, J. (Ed.), *Behavior Genetic Analysis*. New York: McGraw-Hill.
- VANDENBERG, S. G. (Ed.) (1965). *Methods and Goals in Human Behaviour Genetics*. New York: Academic Press.

- VANDENBERG, S. G. (1972). Assortative mating, or who marries whom? *Behav. Genet.*, 2, 127-157.
- VANDENBERG, S. G., MEREDITH, W., and KEESE, A. (1980). Two higher order factors in intelligence test scores. A partial replication of the Cattell-Horn theory of fluid and crystallized intelligence. (In press.)
- VERNON, P. E. (1969). *Intelligence and Cultural Environment*. London: Methuen.
- WILLIAMS, T. (1977). Reply to Professor Lewontin's letter. *Behav. Genet.*, 7, 37.
- WOLIVER, R. (1979). Assortive mating correlations of American and Japanese honeymoon couples, on the MAT. Ph.D. Dissertation, University of Hawaii Library, Honolulu.

(Manuscript received 18th July, 1979)

AN ANALYSIS OF GROUP INTERACTION AND MATHEMATICAL ERRORS IN HETEROGENEOUS ABILITY GROUPS

By NOREEN M. WEBB

(*University of California, Los Angeles*)

SUMMARY. This study investigated group interaction and mathematical errors of 11th-grade students working individually and in four-person heterogeneous ability groups. Averaging over all types of errors, high-ability students did best in the individual condition, low-ability students did best in the group condition, and medium-ability students did equally well in both conditions. When types of errors were distinguished, for high-ability students, the group condition was detrimental to learning new material (the mathematical algorithm) but was advantageous for performance on previously learned material (computational and algebraic manipulations); for medium-ability students, the group condition was good for learning the algorithm but not for performance on computational and algebraic manipulations; for low-ability students, group work was beneficial for performance of new and previously learned material. Group interaction was related to achievement: students who described or received explanations about the mathematical algorithm did well on the algorithm on the test, and students who explained how to perform computational or algebraic manipulations did well on those manipulations on the test.

INTRODUCTION

RECENT research comparing achievement in groups heterogeneous in ability with achievement in individual settings suggests that heterogeneous groups may be beneficial for students at most ability levels. This research does not, however, distinguish between learning of new skills and performance on previously learned skills. The purpose of the present paper is to investigate the influence of the learning setting (heterogeneous ability group or individual) on achievement of new and previously learned mathematical material. The paper not only discusses achievement, but analyses interaction among group members to explain the results of achievement in the group setting.

Nearly all of the research comparing heterogeneous ability group and individual conditions reports benefits of the group setting for students at some ability levels. Results of the studies differ, however, on *which* students benefit from group interaction. Beane and Lemke (1971) and the many studies of tutoring (Cloward, 1967; Lippitt and Lippitt, 1968; Mollod, 1970; Klosterman, 1970; Devin-Sheehan *et al.*, 1976; Smith, 1977), have found that working in heterogeneous ability groups is beneficial for all students. Peterson and Janicki (1979) found that heterogeneous ability groups were advantageous for high-ability students but not for low-ability students, whereas Amaria *et al.* (1969) found that heterogeneous ability groups were advantageous only for low-ability students. Webb (1977, 1980) found that heterogeneous ability groups were advantageous for high-ability and low-ability students, but were detrimental for medium-ability students.

Some of the above investigators advanced 'participation' hypotheses to explain the results of their studies. Amaria *et al.* (1969), Peterson and Janicki (1979), and Webb (1977, 1980) hypothesised teacher-learner relationships among high-ability and low-ability students. Only Peterson and Janicki and Webb investigated interaction among group members to substantiate the hypothesis. Peterson and Janicki found that high-ability students benefited by explaining to other students. In addition, Webb found that low-ability students benefited from being the target of explanations.

In Webb's work, medium-ability students did not benefit from the group experience because they did not actively participate in explaining or soliciting explanations.

Webb's (1980) analysis of group interaction shows that merely participating in a teacher-learner relationship was not sufficient for high achievement. In that study the *content* of explanations given or received was related to achievement on that content. Students who did not explain or receive explanations on a specific task did not perform well on that task on post-tests, regardless of the extent of their participation on other tasks.

Although the analysis of interaction related to specific content helps to clarify results of achievement, none of the research described above has differentiated between students who participate in group interaction only for material they already know and students who participate to help solve new problems. This distinction may help explain the inconsistent results across studies. If, for example, high-ability students in the Amaria *et al.* (1969) study did not participate in solving new problems but only explained what they already knew, then the group experience would not be expected to increase their learning. If, in the Beane and Lemke (1971) study, high-ability and low-ability students participated in group interaction when learning new material, the increased performance for all students would be expected.

The present paper analyses group interaction and achievement related to new and previously learned mathematical material.

METHOD

Sample

The five groups of four students discussed in this paper were part of a larger sample of 66 students who participated in the Webb (1977) study. (The sample and procedures are described in detail by Webb, 1977; only a summary is presented here.) The 66 students were selected from a pool of 181 eleventh-grade students from three high schools in a suburban area in Northern California. On the basis of the students' scores on a battery of aptitude and achievement tests, three general ability strata were defined: high, medium, and low. (The mean IQs of students in the three ability strata were 124, 117, and 105.) Twenty students were assigned to mixed ability groups of four students each. In each group there were two medium-ability students, one high-ability student, and one low-ability student.

Materials and tests

In the Webb (1977) study, students performed three tasks. Two of the three tasks proved to be of equal difficulty and are, therefore, the only tasks reported here. Task I was based on mathematical probability. Each student received instruction on calculating the expected value of a game of chance. Task II dealt with positive and negative number bases. Students were taught to convert numbers presented in different number bases to other number bases. Example tasks and solutions are given in Table 1.

There were two phases of work on a task before testing: training and problem practice. During training, students received instruction on component concepts and skills of a task. The training materials consisted of individual instruction booklets presenting definitions, descriptions, sample exercises and solutions, and exercises for the student to do. During problem practice, students used the components to solve complex problems. A complex problem incorporated all concepts and skills in the training booklet. Accompanying each practice problem was a booklet of hints consisting of a step-by-step solution of the problem, to enable every student to solve the problems.

For each task there was a test consisting of one problem. The problem had the

TABLE 1
EXAMPLE TASKS AND SOLUTIONS

Task	Problem Description	Solution																
I.	Suppose you play a gambling game at a Las Vegas Casino. In this game you have to do two things. First you have to spin a pointer on a spinner that has four colours: red, blue, yellow and green. Each colour is equally likely. Second you have to toss an honest six-sided die. You win \$50 if the pointer lands on red <i>or</i> if the die comes up showing 4. If anything else happens you don't win anything. What is the expected value of this game?	Expected value = $= 50 [P(\text{red or } 4)]$ $= 50 [P(\text{red}) + P(4) - P(\text{red and } 4)]$ $= 50 \left[\left(\frac{1}{4}\right) + \left(\frac{1}{6}\right) - \left(\frac{1}{24}\right) \right]$ $= 50 \left(\frac{5}{6}\right)$ $= \$18.75$																
II.	You are on an interplanetary mission aboard the Starship Enterprise. Your role is that of a Federation accountant and diplomat. Your mission is to collect money from Omega I, II and III to give to Alpha I and Alpha II to help pay for damages from the recent planetary war. Your accounting skills are needed because the monetary system of each planet is in a different base. You are given the following information: Omega I is willing to give 211_3 units to Alpha I and 211_3 units to Alpha II. Omega II is willing to give 52_{10} units to Alpha I and 52_{10} units to Alpha II. Omega III is willing to give 333_{-4} units to Alpha I and 333_{-4} units to Alpha II. Alpha I's monetary system is in base 6. Alpha II's monetary system is in base -8. How much will each planet (Alpha I, Alpha II) receive in its own base?	$211_3 = 2(3^2) + 1(3^1) + 1(3^0) = 22_{10}$ $52_{10} = 52_{10}$ $333_{-4} = 3(-4^2) + 3(-4^1) + 3(-4^0) = 37_{10}$ $\text{Sum} = 22_{10} + 52_{10} + 37_{10} = 111_{10}$ For Omega I (dividing by 6): <table><thead><tr><th>Quotient III</th><th>Remainder</th></tr></thead><tbody><tr><td>18</td><td>3</td></tr><tr><td>3</td><td>0</td></tr><tr><td>0</td><td>3</td></tr></tbody></table> So $III_{10} = 303_6$ For Omega II (dividing by -8): <table><thead><tr><th>Quotient III</th><th>Remainder</th></tr></thead><tbody><tr><td>-13</td><td>7</td></tr><tr><td>2</td><td>3</td></tr><tr><td>0</td><td>2</td></tr></tbody></table> So $III_{10} = 327_{-8}$	Quotient III	Remainder	18	3	3	0	0	3	Quotient III	Remainder	-13	7	2	3	0	2
Quotient III	Remainder																	
18	3																	
3	0																	
0	3																	
Quotient III	Remainder																	
-13	7																	
2	3																	
0	2																	

same format as the problems solved during problem practice. All test problems were scored on a 0- to 20-point scale. Credit was given for steps performed correctly.

Procedure

One task was assigned for each session. Sessions for Task I and Task II were spaced one week apart.

Training. The procedure during training was identical for all students. The experimenter explained the purpose and nature of the study and procedures. Students worked alone on the training booklets and checked their answers to the exercises with the experimenter. Time limits for training were set to allow all students to finish the training booklets. The training booklets included enough additional exercises to occupy all time allotted to training.

Problem practice. Problem practice was done individually for Task I and in interacting groups for Task II. During problem practice for Task I, all students worked on the practice problems alone. Students were instructed to use the booklets of hints to help them solve the problems and to consult the experimenter only after studying all hints.

During problem practice for Task II, all students in a group sat at one table. Each group was told not to divide the work, but to work as a four-person group to solve each practice problem. Students were instructed to ask questions of teammates and to explain how to solve the problems to any group member who was confused.

What group members said during group problem practice was recorded on the audio recorder. As for Task I, students had available the booklets of hints.

Testing. After problem practice for Tasks I and II, students returned to their original tables to do the test problems individually. Students were not permitted to use hints or other materials, or to consult the experimenter or other students. To relieve pressure on low-ability students, there was no time limit for any test.

Error types

The errors made on the tests were classified as (1) errors specific to the algorithm for carrying out the task (new material), and (2) computational or algebraic errors not specific to the algorithm (previously learned material).

(1) *Algorithmic error.* A student set up the algorithm incorrectly or carried out a step of the algorithm incorrectly. An example of an algorithmic error for Task I (probability) is:

The probability of Event A and Event B = $P(\text{Event A}) + P(\text{Event B})$.

Error: Incorrect rule for computing combined probabilities.

Correct answer: $P(\text{Event A}) \times P(\text{Event B})$.

An example of an algorithmic error for Task II (negative number bases) is:

$$212_{-4} = 2(2)^{-4} + 1(1)^{-4} + 2(0)^{-4}.$$

Error: Incorrect rule for converting a number in base negative 4 to a number in base 10 (base and exponent were switched).

Correct answer: $2(-4)^2 + 1(-4)^1 + 2(-4)^0$.

Post-experimental questionnaires and interviews revealed that the algorithms in Tasks I and II were new to the students.

(2) *Computational or algebraic error.* A student made an error of computation or algebra that was unrelated to the steps required in setting up or carrying out the algorithm. An example of a computational or algebraic error for Task I is:

$$59/416 = 7.05.$$

Error: Quotient and divisor were switched.

Correct answer: 0.017.

An example of a computational or algebraic error for Task II is:

$$(3)^2 = 6.$$

Error: Base was multiplied by 2 instead of being squared.

Correct answer: 9.

Nearly all of the computational and algebraic errors seemed to be the result of carelessness rather than lack of understanding or knowledge. Results of a questionnaire asking students about their previous exposure to the topics covered indicated that students had previously received instruction on basic principles of computation and algebra. When confronted with their errors of computation or algebra, nearly all students remarked that they were careless, that they 'knew better.'

Group interaction variables

Based on previous studies of learning in groups (Webb, 1977, 1980; Peterson and Janicki, 1979), the following observation categories were expected to relate to algorithmic errors: (1) stating or describing an algorithm (not in response to a specific student's need), (2) explaining an algorithm in response to a specific student's need, and (3) receiving an explanation of an algorithm. The following categories were expected to relate to computational or algebraic errors: (4) carrying out a computational or algebraic manipulation (not in response to a specific student's need), (5)

explaining how to carry out a computational or algebraic manipulation in response to a specific student's need, and (6) receiving an explanation of a computational or algebraic manipulation. 'Responding to student need' was defined as responding to a student's question, statement of confusion, or error. From transcripts of the audio recordings, the number of instances of each of the above categories of behaviour was coded for each student.

RESULTS

The results will be presented in four parts. First, performance in the individual and group settings will be described without distinguishing between error types for each ability level. Second, error type will be taken into account in the comparison of performance in individual and group settings. Third, group interaction will be described for each ability level. Finally, results of group interaction and achievement will be synthesised to help explain differences in performance across ability levels and work settings.

Total test performance

Table 2 summarises the incidence and severity of errors on tests in the individual and group settings. The bottom row in each section of the table shows the incidence and severity of errors without differentiating by type of error. As can be seen in the last two columns of Table 2, average performance was about the same on tests in the individual and group settings. Over half of the students made errors. Those who made errors missed, on the average, about a quarter of the points on each test.

Although average performance was the same on tests in the individual and group settings, performance at each ability level differed across learning settings. In the individual condition, as would be expected, high-ability students did best and low-ability students did worst. The group setting tended, however, to equalise the performance of students at all ability levels. The same proportion of students at each ability level made errors, although the errors made by lower-ability students were more severe than those made by higher-ability students. To compare learning settings, high-ability students did worse learning in groups than they did learning individually. Medium-ability students performed the same in groups as they had in the individual setting. Low-ability students working in groups improved markedly over their performance in the individual setting.

When types of errors are not distinguished, then, working in groups, compared to learning individually, was detrimental for high-ability students, was neither beneficial nor detrimental for medium-ability students, and was markedly beneficial for low-ability students.

Performance by error type

Examination of Table 2 shows that the results for total test performance mask results for each error type. Whereas the findings for the total test apply to low-ability students when error types are distinguished, the conclusions about performance for high-ability and medium-ability students depend upon the type of error. High-ability students made more errors carrying out the algorithm (algorithmic errors) when working in groups than when working individually, but made fewer computational or algebraic errors. Medium-ability students, on the other hand, made fewer algorithmic errors when working in groups than when working individually, but made more computational and algebraic errors.

Thus, whether working in groups is interpreted as beneficial or detrimental compared to individual work depends upon the type of error that is emphasised. Working in groups does not seem to help high-ability students learn new material

TABLE 2
INCIDENCE AND SEVERITY OF ERRORS IN INDIVIDUAL AND GROUP SETTINGS

Error type	High-ability		Medium-ability		Low-ability		All students	
	% students	% test points lost*	% students	% test points lost*	% students	% test points lost*	% students	% test points lost*
Individual Setting								
Algorithmic	0	0	50	23	100	43	50	33
Computational or algebraic	20	5	30	8	60	8	35	8
All errors combined	20	5	60	23	100	48	60	32
Group Setting								
Algorithmic	60	11	30	26	40	48	40	26
Computational or algebraic	0	0	60	6	20	5	35	6
All errors combined	60	11	60	19	60	33	60	21

* % test points lost per student who made errors.

Note: Because the number of students differs across ability levels, percentages are used to allow direct comparisons across categories.

(the algorithms), but seems to make them more careful when doing work they already know (computations and algebraic manipulations.) Working in groups seems to help medium-ability students learn new material, but seems to make them careless about previously learned operations or definitions. Group work is good for low-ability students on all counts. We turn now to an intensive examination of group process that may explain the differences in performance across error types.

Group interaction

Analysis of the roles that students played and their experiences in the groups help to explain the performance by error type at each ability level. Group interaction is analysed separately for content related to (1) the algorithms, and (2) computational and algebraic manipulations.

High-ability. Table 3 summarises the results of the group interaction analysis. As can be seen in Table 3, high-ability students seldom described or explained specific features of the algorithms and never received explanations of the algorithms. Transcripts of group work showed that, in nearly every group, high-ability students directed the group work by stating overall objectives and delegating work to other members of the group. The following excerpt is an excellent example of a high-ability student describing the objectives of the task, without providing any clue about how to carry it out.

"Each planet is giving an equal amount to either planet so you just . . . figure out what it would be for each of those. And that'll just tell you what's given to Alpha I and what's given to Alpha II. It's like you have to figure out what this is in base ten and that is in base ten; that's already in base ten. Then you have to transpose the base ten thing into base six."

Or, as the high-ability member of another group stated more succinctly, "First translate this to base ten, and then translate it to base six." Delegations of responsibility included "Write out the quotient thing" and "Now add those."

High-ability students spent much time calculating and explaining how to perform calculations, but received few explanations about calculations. High-ability students often checked and corrected the calculations of other members: "Three squared is nine times two is 18"; "It's negative one which is 34. Negative 14 times a remainder of one. That goes four"; and "No, because two times six squared is not the same as two times three squared." Some high-ability students even explained the steps in a specific calculation: "Since it's negative 52 we've got to divide it in positive seven times to get 56 and then subtract before we get negative four. I mean positive four."

TABLE 3
MEAN NUMBERS OF INSTANCES OF GROUP INTERACTION VARIABLES

Group Interaction Variable	Ability		
	High (N = 5)	Med (N = 10)	Low (N = 5)
Algorithm			
Stating or describing algorithm	2.4	4.6	1.4
Explaining algorithm in response to need	2.2	2.6	1.2
Receiving explanations of algorithm	0.0	1.2	6.2
Computational or algebraic manipulations			
Stating or describing calculations	23.4	15.7	10.6
Explaining calculations in response to need	8.8	2.4	1.0
Receiving explanations of calculations	1.0	2.6	8.4

High-ability students, then, spent more time calculating and explaining how to perform the calculations than describing or explaining the algorithms.

Medium-ability. Whereas high-ability students rarely described the algorithm, medium-ability students not only set up the algorithm and carried out the steps, but often explained how to carry out the algorithm to other members of the group. They rarely received explanations, however. When describing the algorithm, medium-ability students usually substituted the appropriate numbers: "Three times negative four squared plus three times negative four to the one plus three times negative four to the zero." Some descriptions were quite detailed. The following excerpt shows a medium-ability student carrying out the algorithm after the high-ability student stated the objective of the problem.

HIGH: Once you get that number you have to change it into base six.

MEDIUM: To do it, we just divide it [421] by six and that'll give us seventy remainder of one. Divide by six again and we get one, which would be eleven, remainder of four. Then we divide it by six again and get one with a remainder of five. Divided by six again, we get zero with a remainder of one. So we have 1541 base six.

Medium-ability students participated in calculations, but rarely explained how to carry them out and rarely received explanations about them. Medium-ability students often failed to detect errors made by other members. No medium-ability student detected the following error made by another group member: "Seven squared times six. What's 14 times six?" Some medium-ability students expressed lack of confidence in their ability to perform calculations or asked other group members for the answers to simple calculations: "Okay, so then we have, what's 16 times three?"; "Six times 50 is going to be 300. Right?"; "All we did is that negative times negative will give us . . . plus . . . I think maybe that might be wrong"; "39 plus . . . Go ahead, I can't do it." Some medium-ability students did not even attempt calculations. In summary, although medium-ability students typically described the algorithm, they often seemed unsure of their calculations.

Low-ability. Low-ability students rarely participated in setting up or explaining the algorithms, but they often solicited and received explanations from other members. The following excerpt is typical of the explanations that low-ability students received:

LOW: So what did you do here?

MEDIUM: I changed that to a negative base, this to base 10. It's the same thing here. Two times negative five squared plus zero times negative five to the first power, plus four times negative five to the zero power.

LOW: So okay, that's 25.

MEDIUM: Times two is 50. And that would be zero 'cause zero times anything is zero, and that's four times one 'cause anything to the zero power is one.

LOW: Yeah.

MEDIUM: Now we're going to add together all that it [Alpha] will get. It [Alpha] gets that [points] and that [points] and that [points]. We changed those into base 10. So we're going to change it [the total] to its right base.

LOW: Oh, okay, then we have to change it to base six.

In explanations to low-ability students, other members made certain that the low-ability member understood how to carry out the algorithms.

Whereas low-ability students rarely volunteered to set up the algorithm, they did perform some calculations. They often received explanations about how to perform the calculations, particularly when they made errors. For example, in response to a low-ability student who added a positive 48, a negative 12, and a positive 3 and obtained the sum of 63, another student said, "No, you *subtract* 12 from 48 before

adding three", rather than merely stating the correct answer, which was the response typically given to other students who made computational errors. Low-ability students, then, often received explanations about the algorithm and about computational and algebraic calculations.

Summary of group interaction and ability. As the previous discussion suggests, ability and group process variables were highly correlated. Higher-ability students gave descriptions and explanations more often than lower-ability students (correlations ranged from 0.39 to 0.78, all significant at $P < 0.04$). Lower-ability students received explanations more often than higher-ability students (correlations ranged from 0.54 to 0.71, all significant at $P < 0.01$).

Group process and achievement

To assess the relation between group process and achievement independent of the effects of ability, partial correlations were computed between group process variables and achievement, controlling for ability. The measure of ability was the test composite used to classify students into ability strata before assignment to groups.

Stating or describing the algorithm and receiving explanations about the algorithm were positively related to the number of test points correct for the algorithm. (The partial correlations were 0.37 ($P < 0.06$) and 0.63 ($P < 0.00$), respectively.) Giving explanations about the algorithm in response to student need did not relate to achievement. For calculations, only explaining how to perform computational or algebraic manipulations in response to student need was related to the number of test points correct for calculations. (The partial correlation was 0.39, $P < 0.05$). Neither describing the algorithm nor receiving explanations about the calculations related to achievement.

DISCUSSION

The results of analyses of achievement and group interaction reveal several important findings. First, averaging over ability levels, students performed equally well on achievement tests in the group setting and in the individual setting. Second, the effect of the setting differed depending upon the ability level of the student. High-ability students did better working individually; medium-ability students did equally well in both settings; and low-ability students did best working in groups. Third, there was an interaction between learning setting, ability, and type of performance (algorithmic or computational and algebraic). For high-ability students, heterogeneous ability groups were advantageous for performance of computational or algebraic manipulations, but not for learning the algorithm. For medium-ability students this finding was reversed. For low-ability students, group work was beneficial in both areas.

Most importantly, the experience of students in the group was related to their test performance. Students who gave many descriptions of the algorithm, received many explanations of the algorithm, or explained how to perform computational or algebraic manipulations did better than students who displayed less of the above behaviour. Medium-ability students tended to give the most descriptions of the algorithm; low-ability students received the most explanations of the algorithm; and high-ability students gave the most explanations of computational and algebraic manipulations.

Describing the algorithm may have helped students to understand and to solidify in memory how to carry out the algorithm. Students may be better able to locate their own area of difficulty by describing it than by listening to others describe it. Being the target of explanations about the algorithms may help students to understand the algorithm more than merely reading text materials or listening to descriptions. The

other students in the group may understand a particular member's difficulty and be able to provide descriptions of the troubling materials in terms that the student experiencing the difficulty can understand. Group members often remarked that the explanations they received helped them understand the task better. Providing explanations of computational and algebraic manipulations in response to a specific student's need may focus the explainer's attention on particularly difficult calculations. By attending to manipulations troublesome to others, the explainer may take more care when carrying out those particular manipulations on tests.

The ineffectiveness (for the explainer) of giving explanations in response to student need was unexpected. The low correlation may be due to restriction of range; compared to other categories of behaviour, students gave few explanations of the algorithm in response to need. The lack of relation between describing the calculations and performance was not unexpected. Students tended to perform calculations by rote. The frequency of careless errors in the group work suggests that students performing calculations were attending more to the algorithms than to the calculations. Another student's call for help was needed for students to focus on the calculations. Accounting for the lack of relation between receiving explanations about calculations and performance, students' difficulty in carrying out computational or algebraic manipulation seemed to indicate deficits in those skills. The group sessions were too brief to remedy deficiencies in computation or algebra that months or years of classroom instruction had not remediated.

The findings of achievement and group interaction reported here may help explain the inconsistent results of studies comparing achievement in different learning settings. For example, Peterson and Janicki (1979) found that receiving explanations was not related to achievement. The lack of relation may be due to the content of the explanations. If students received explanations on material they already understood, rather than on the new material, then receiving explanations would not be expected to relate to achievement on the new material.

Clearly, studies of group learning should include information on group interaction and achievement related to specific content areas of the material being learned. Any conclusions on the advantages of a group setting are linked to specific experiences of each group member. Further research is needed to determine whether the specific roles in group interaction found in this study will occur for persons of different ages and for groups learning different subject matters. In regard to practical applications of grouping, educators using ability groups in classrooms need to monitor group interaction to understand the impact on achievement and to encourage group experiences that are beneficial for achieving desired learning objectives.

REFERENCES

- AMARIA, R. P., BIRAN, L. A., and LEITH, G. O. M. (1969). Individual vs. co-operative learning: influence of intelligence and sex. *Educ. Res.*, 11, 95-103.
- BEANE, W. E., and LEMKE, E. A. (1971). Group variables influencing the transfer of conceptual behavior. *J. educ. Psychol.*, 62, 215-218.
- CLOWARD, R. D. (1967). Studies in tutoring. *J. exper. Educ.*, 36, 14-25.
- DEVIN-SHEEHAN, L., FELDMAN, R. S. and ALLEN, V. L. (1976). Research on children tutoring children: critical review. *Rev. educ. Res.*, 46, 355-385.
- KLOSTERMAN, R. (1970). The effectiveness of a diagnostically structured reading program. *Reading Teacher*, 24, 159-162.
- LIPPITT, R., and LIPPITT, P. (1968). Cross-age helpers. *Today's Education*, 57, 24-26.
- MOLLOD, R. W. (1970). Pupil tutoring as part of reading instruction in the elementary grades. (Doctoral dissertation, Columbia University.) *Dissertation Abst. Int.*, 31, 2260B. (University Microfilms No. 70-18, 835).
- PETERSON, P. L., and JANICKI, T. C. (1979). Individual characteristics and children's learning in large-group and small-group approaches. *J. educ. Psychol.*, 71, 677-687.

- SMITH, C. C. (1977). Partner learning: peer tutoring can help individualisation. *Educational Leadership*, 34, 361-363.
- WEBB, N. M. (1977). *Learning in individual and small group settings*. (Technical Report No. 7). Stanford, Calif.: Aptitude Research Project, School of Education, Stanford University.
- WEBB, N. M. (1980). Group process and learning in an interacting group. *Newsletter of the Laboratory of the Comparative Human Cognition*, 2, 10-15.

(Manuscript received 4th February, 1980)

SHERLOCK HOLMES MEETS OTHELLO: A MDS ANALYSIS OF LITERARY CHARACTERS

BY G. W. RUSSELL AND W. B. LAMBERT

(Department of Psychology and Department of English, University of Lethbridge,
Alberta, Canada)

SUMMARY. Changes in students' perceptual organisation of the characters from *Othello* after three weeks of study and lectures were assessed using multidimensional scaling procedures. Senior Sherlockians also provided dissimilarity ratings of a sample of characters from the Conan Doyle stories. *Evaluative* and *Sophistication* dimensions dominated the solutions for both groups of judges. The students used an important *Demeanour* dimension and a *Sexuality* dimension in the pre-lecture stage. Discussion centred on the emphases and interpretations offered in the lectures on *Othello* and the extent to which they may have effected changes in the order and distances of major characters along the principal dimensions.

INTRODUCTION

A LONG-STANDING problem of some perplexity for educators has been the measurement of a reader's response to literature. The point of contention between those who would attempt to operationalise such wispy concepts as appreciation, understanding, etc. and those who would throw up their hands in utter horror at such efforts has been cogently summarised by Forehand (1966): "Educational objectives phrased in terms of understanding, appreciation, enrichment, and the like are not meaningful unless they can be translated into observations of the behaviour of students. Until we can obtain objective, reliable measures of what the student does, the terms will reverberate endlessly among the private meanings that different speakers give them." A teacher of English might reply: "But measuring instruments, in the attempt to capture a response to literature, in fact destroy it" (p. 369). Regardless of our predilections, measurement procedures and research strategies are being developed and applied on a broad scale to artistic creations in music, art and literature (e.g., Berlyne, 1974; Lindauer, 1974). Whether traditionalists see experimental aestheticians as developing a valuable approach to aesthetics or, in Berlyne's (1976) words: "as barbarian despoilers of their territory" (p. 86) is a debate of little moment. Judgments about their ultimate value to literature and education are perhaps best left to those 'state of the arts' reviews still to be written.

In the introduction to his investigation of the responses by college freshmen to three American novels, Wilson (1966) observed that classroom study of a novel clearly changes the way students respond to it but there is a dearth of evidence attesting to the nature of those changes. Analysing student protocols, he found an increase in the number of interpretational responses with an attendant decrease in the number of literary judgments from a pre- to a post-discussion stage. Reactions scored as interpretational dealt with attempts "to discover the meaning of the story, the motivational forces, the nature of the characters" (p. 9) whereas literary judgments included responses "concerned with the author and his method, plot mechanisms, adequacy of characterisation, language, style, and general evaluations of literary worth" (p. 8).

Though its applications to literature have been few, one potentially valuable investigative technique is multidimensional scaling (MDS). Romney *et al.* (1972) provide a discussion of the limitations and applications of the procedures to diverse subject areas while a paper by MacCallum (1974) describes their relationships with

factor analysis. It would be an extravagant claim, indeed, which suggested that all aspects of the literary experience could be revealed through the use of MDS procedures. However, it is suggested that they can serve an important function in identifying those dimensions used by students in processing information from the complex fabric of human interactions which constitute the focus of so many works. Typically, the procedures call for the subject to rate the degree of (dis)similarity between each of all possible pairs of a set of n objects. The subjects are free to consciously or intuitively use any bases they choose in making their similarity judgments. The resulting analysis provides a space—usually Euclidian—in which the n points are so arranged as to correspond to the set of n objects, with the interpoint distances representing the judged similarity between pairs of the objects. The final geometric configuration is commonly taken to be a representation of the subjects' perceptual organisation of the domain from which the n objects were originally chosen. Thus, the approach offers insights into the structure of actual vantage points which subjects use in attempting to understand an aesthetic work.

The usefulness of MDS procedures in identifying the nature and number of dimensions used by students in considering the works of an author has been established by Wainer and Berg (1972). Senior French majors used *Violence* and *Fulfillment* dimensions in their perception of nine stories by de Maupassant. Neither dimension was related to the students' liking of the stories. Another application of MDS procedures to literature was undertaken by Rosenberg and Jones (1972) in their investigation of the dimensional perspectives of Theodore Dreiser utilising the descriptive traits he applied to his female characters in *A Gallery of Women*. The most parsimonious arrangement of Dreiser's view was a representation in four dimensions: (i) sex of character, (ii) conformity, (iii) the pervasive evaluative dimension, and (iv) potency, a solution described by the authors as largely consistent with Dreiser's personal history and involvement with women. The importance of a writer's view notwithstanding, a related concern is *changes* effected in the perceptual organisation of the reader through study of an aesthetic work. The suggestion has been made earlier (Klein, 1968) that the use of MDS may adequately reflect changes in students' judgments about art works following a course in drawing or art appreciation. Whether the stimuli are drawn from the domains of art, music or literature, the task facing the investigator is to demonstrate convincingly that any changes in perceptual organisation arose from the reading or study experience rather than any number of sources of influence present between the pre- and post-test stages of a design (e.g., history, maturation). To date this concern has been met retrospectively by drawing attention to the parallels between changes in the configuration and those changes the writer feels would likely follow from in-depth study and the presentation of professional perspectives. Several MDS studies have attempted to examine *changes* in perceptual structure over time and to link those changes to intervening events. In an organisational setting, Jones and Young (1972) observed meaningful structural changes over two years in the self-perception of a cohesive group of laboratory workers. Students used the same three dimensions over a ten-week Developmental Psychology course in their perceptions of 16 course-related concepts (Wainer and Kaye, 1974). Of special significance was a pronounced shift in student position towards the position of the instructor within the spatial configuration. More recently, Berg and Wainer (1976) had students scale translations of nine short poems by Baudelaire, a 19th-century French poet, before and after two weeks of classroom discussion. The same single dimension offered the best solution on both occasions though the order and distances on that dimension changed somewhat, apparently again in ways consistent with the course of classroom discussions. The same poems, untranslated from the original French, were also scaled by different students in a French literature class. Curiously, the best solution was a representation in *two* dimensions, both before and after the instruction period.

Applications of MDS procedures to literature have so far been limited to the use of descriptive traits (Rosenberg and Jones, 1972), poems (Berg and Wainer, 1976) and short stories (Wainer and Berg, 1972) as stimuli. None has dealt with the reader's perceptual organisation of characters within a work, nor, for that matter, with changes in perceptual organisation from one's initial, superficial familiarity to that at a stage following intensive study. Suffice it to say, there is no reason to assume that similar bases for judgments would be used where characters rather than poems or stories themselves constitute the stimuli to be scaled. In attempting to elaborate the nomological network (Cronbach and Meehl, 1967) in which the construct of literary dimensionality occurs, qualitatively different types of literature were used as were judges differing in age, cultural experience, and expertise. The present investigation, in pursuit of these issues, used MDS procedures in a before-after design with Shakespeare's *Othello*, and in a post-test only condition with Conan Doyle's engaging Sherlock Holmes adventures.

METHOD

The subjects were mostly first year students ($N = 37$, mean age 19.7 years, SD 3.4) enrolled in an Introduction to Literature course at the University of Lethbridge. *Othello* had been assigned as a reading prior to its being analysed in lecture form by the second author. All students (24 males and 13 females) scored successfully (at least 12/15 on a multiple choice quiz measuring their reading familiarity with the principal characters and their relationships to each other (e.g., "Desdemona is suspected of adultery with (Othello, Iago, Montano, Brabantio, Cassio). She really engages in adultery with (Othello, Iago, Montano, Brabantio, Cassio, none of these).") At this point the students were asked for their co-operation in the present study, described as an interdisciplinary scaling project being undertaken by members of the English and Psychology departments and totally unrelated to the evaluation of their performance in the course. There was general enthusiasm for the task and none chose the option of leaving the scales blank. Subjects then provided ratings of dissimilarity between all possible pairs of the 13 figures (the *dramatis personae*) of the work along accompanying seven-point scales. No mention was made of the fact that the same subjects would be asked to repeat their ratings three weeks later.

During the interval students were subjected to approximately nine hours of lectures, attendance at which exceeded 95 per cent. Given the introductory nature of the course, the lectures were eclectic, drawing upon the three common modes of literary study, viz. formalist, sociocultural, and psychoanalytic. The lectures emphasised the dramatic complexity which forms the structure of the play around interpersonal relationships, manipulation, jealousy, power-lust, and the difficulty of retaining control and preserving order in such volatile circumstances. Among other things, it was stressed that the complex poetic imagery of the play tends to reinforce our understanding of the changing relationships of the characters in the general deterioration of political and personal order. For example, in the earlier parts of the play, Iago's speeches are laden with animal images used to describe human relationships. Somewhere in the middle of the play, Othello begins to use such images, and by the end he has reduced all human relationships to animal terms (Heilman, 1956, p. 104 ff).

After providing the second set of dissimilarity ratings, subjects were then asked to make ratings of each character along a set of 11, seven-point bipolar adjective scales. The dissimilarity matrices served as input to TORSCA 9 (Young, 1968), a program which rotated the solution to 'simple structure' by means of a varimax routine. It should be noted that at this point the final examination covering *Othello* and other literary works was still three weeks distant.

In meeting the requirement of a qualitatively different type of literature, we also had to secure judges who were 'expert' on the work(s) chosen. For this reason we enlisted the support of members of the Sherlock Holmes Society of London ($N = 18$, mean age 50.6 years, $SD 20$), expert by virtue of their long-standing membership in the society. They, too, made dissimilarity judgments between all possible pairs of an arbitrary sample of the major figures in the Sherlock Holmes adventures. Sir Arthur Conan Doyle was included as a marker variable along with 13 of his major characters. As in the former analysis, supplemental ratings were intended to provide a measure of empirical verification for otherwise subjective interpretations given dimensions arising from the analyses.

RESULTS

Interjudge consistency was determined by a procedure proposed by Winer (1971). Analyses of variance were performed for both sets of dissimilarity judgments with F values corresponding to reliability coefficients of 0.88 for *Othello I* and 0.87 for *Othello II*. The overall reliability of the supplemental ratings was 0.96. Scaling of the initial *Othello I* and the later *Othello II* dissimilarity matrices was done in two to five dimensions. Stress values were 11.7 per cent, 8.0 per cent, 5.2 per cent, 2.6 per cent, and 11.3 per cent, 6.4 per cent, 4.0 per cent, 2.0 per cent, respectively. Four dimensional solutions seemed optimal in the light of criteria developed by Wagenaar and Padmos (1971) for interpreting the significance of 'stress' values and the overriding consideration of interpretability. The correlations between the dimensional loadings and supplemental ratings improved the clarity of most dimensional interpretations, a hazardous task even so.

The interpretations given the rotated *Othello I* dimensions included: (i) an *Evaluative* dimension on which the ill-intentioned Iago and jealous Roderigo anchored one end in opposition to the virtuous Desdemona and Cassio; (ii) *Demeanour* in which the lighthearted Bianca, the Clown, and in some respects Iago, cluster at one end of the dimension and the more sombre figures of Brabantio, Gratiano and Lodovico, the other; (iii) *Sophistication*, reflecting decreasing degrees of worldliness and social class from the Senator Montano to the Clown; (iv) *Sexuality*, along which females cluster in order of sexuality on one end of the scale with the males ordered on the other in terms of increasing virility and desirability.

Three dimensions of the four-dimensional, post-lecture solution were given the same interpretations, i.e., *Evaluative*, *Demeanour*, and *Sophistication*, two of which are presented in Figure 1.

The earlier *Sexuality* dimension was replaced by a *Control* dimension reflecting increasing degrees of legitimate social and political authority. Thus, Roderigo and Iago are essentially devoid of power, seeking merely to usurp the rightful authority held or delegated by the Duke and Montano among others. The correlations presented in Table 2 between the rating scales and dimensional loadings were especially helpful in identifying the nature of the principal dimension, somewhat less so with the remaining interpretations.

A reliability coefficient of 0.85 indicated a satisfactory degree of judgmental consistency among the Sherlockians. The reliability of the supplemental rating scales was 0.92 while the stress values associated with solutions in two to five dimensions were 11.9 per cent, 7.1 per cent, 5.2 per cent, and 3.2 per cent.

A three-dimensional solution (Stress = 7.1 per cent, dare we say The Seven Percent Solution?) was found to underlie the perception of characters from the Conan Doyle stories with the pervasive *Evaluative* dimension again dominating the structure. The virtuous Mrs. Hudson, Dr. Watson, and Inspector Gregson were clustered at one end of the continuum with the evil Dr. Grimesby Roylott, Mr. Stapleton, Bodymaster

TABLE 1
VARIMAX ROTATED CONFIGURATION (OTHELLO I AND II)^a

	I Evaluative	II Demeanour	III Sophistication	IV Sexuality	IV' Control
Desdemona	34 ^b 30 ^c	02 -22	13 49	10	-18
Cassio	32 23	03 -12	08 27	-04	-13
Duke of Venice	27 41	18 10	07 12	-22	-32
Gratiano	13 -03	28 20	10 37	08	10
Emilia	12 08	-11 -32	03 10	34	12
Montano	11 01	-03 -01	30 -05	-08	-42
Othello	09 50	-05 -02	11 -01	-36	10
Lodovico	03 -01	23 26	20 16	05	-29
Brabantio	-03 -03	29 41	-11 -03	-07	11
Bianca	-09 -04	-52 -52	-12 -24	13	13
Clown	-11 -12	-20 -03	-54 -61	07	04
Roderigo	-45 -39	03 -06	-31 -16	-06	46
Iago	-72 -90	-16 33	05 -40	05	28

^a Decimals omitted.

^b Othello I (Stress = 5.2%, TORSCA 9).

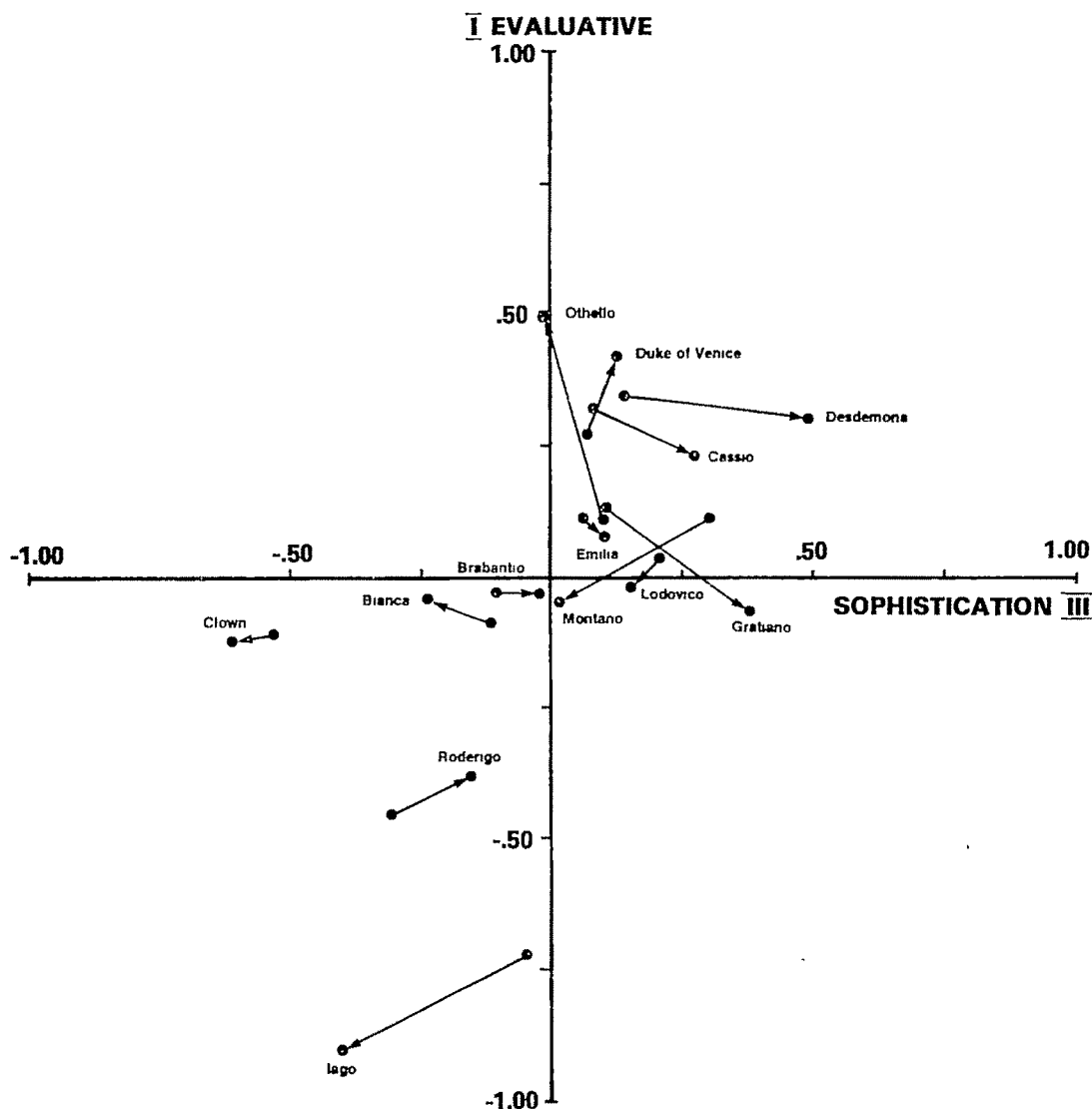
^c Othello II (Stress = 4.0%, TORSCA 9).

McGinty, Professor Moriarty, and Colonel Moran (the second most dangerous man in London) anchoring the other end. This solution did not offer a ready interpretation of the second dimension whereas the third dimension correlated highly with the rating scales of simple-complex, sophisticated-unsophisticated, and imaginative-unimaginative, prompting an interpretation of personal *Sophistication*.

Acting upon a recommendation by Funk *et al.* (1976) arising from evidence that recent non-metric procedures may provide a more stable solution, the original dissimilarity matrix was rerun on KYST (Kruskal *et al.*, 1973), in hopes of resolving the nature of the second dimension. Scaling was done in two to five dimensions with resulting stress values of 10.0 per cent, 5.9 per cent, 3.7 per cent and 2.3 per cent. The KYST solution and its correlations with the rating scales are presented in Tables 3 and 4 respectively. Again, a three-dimensional configuration seemed ideal from the standpoint of interpretability and the negligible improvement in fit to be gained with additional dimensions. The principal dimension was *Evaluative* in nature while the *Sophistication* dimension emerged as second in importance. On a continuum of sophistication Holmes and Irene Adler are seen as more complex, imaginative and flamboyant in contrast with the simple, drab, and uninspiring Inspectors Lestrade and Gregson and their villainous adversaries, Bodymaster McGinty and the King of Bohemia. The third KYST dimension was somewhat more amenable to interpretation and was tentatively given the label of *Activity* primarily because of its 0.67 correlation with the active-passive scale.

The projection of Conan Doyle's characters on the *Evaluative* and *Sophistication* dimensions thus formed three distinct and interpretable clusters. Examination of the

FIGURE 1
PRE- AND POST-LECTURE POSITION OF *Othello* CHARACTERS



clusters presented in Figure 2 reveals a striking degree of consistency with the roles portrayed throughout the stories. The archvillains anchor the negative end of the *Evaluative* continuum being at the same time widely dispersed along the *Sophistication* dimension from the brutish Bodymaster McGinty to Holmes' brilliant adversary, Professor James Moriarty. A second cluster is comprised of goodly, well-intentioned, somewhat inept individuals who in large measure serve as foils to Sherlock Holmes. Both Holmes and Irene Adler are in some ways flawed and thus are evaluated neutrally. However, they and, to a lesser extent, the Baker Street Irregulars are positioned

TABLE 2
CORRELATIONS BETWEEN OTHELLO DIMENSIONS AND RATING SCALES^a

	I Evaluative	II Demeanour	III Sophistication	IV Sexuality	IV Control
Trustworthy-	-.99 ^b	-.28	-.38	.04	
Untrustworthy	-.90 ^c	.33	.75		.65
Good-Evil	-.97	-.27	-.27	.08	
	-.91	.35	.69		.61
Moral-Amoral	-.91	-.51	-.39	.15	
	-.83	.03	.75		.63
Insane-Sane	.85	.27	.29	.29	
	.61	-.27	-.67		-.68
Promiscuous-Chaste	.73	.44	.37	-.16	
	.68	.05	-.63		-.50
Not Jealous-Jealous	-.61	-.19	-.15	-.48	
	-.30	.19	.48		.62
Passive-Active	-.40	-.09	.14	-.44	
	-.25	.29	.16		.37
Uninteresting-Interesting	-.22	-.25	.07	-.23	
	-.09	-.03	.03		.32
Minor-Major	-.10	-.05	.25	-.20	
	-.02	.05	.21		.26
Intelligent-Unintelligent	-.08	-.29	-.39	.29	
	.02	-.56	.10		.41
Weak-Strong	-.02	.19	.43	-.29	
	.57	.53	-.02		-.24

^a Decimals omitted. ^b Othello I. ^c Othello II.

TABLE 3
FINAL CONFIGURATION IN THREE DIMENSIONS (SHERLOCK HOLMES)^a

	I Evaluative	II Sophistication	III Activity
Mrs. Hudson	1.29	-.25	-.41
Dr. Watson	.90	-.19	-.16
Inspector Gregson	.85	-.46	.25
Conan Doyle	.69	.41	-.25
Inspector Lestrade	.65	-.59	.58
Baker Street Irregulars	.45	.44	.60
Sherlock Holmes	.19	.83	-.05
Irene Adler	-.19	.88	-.38
King of Bohemia	-.60	-.53	-.94
Col. Moran	-.73	-.25	.24
Professor Moriarty	-.75	.40	.21
Bodymaster McGinty	-.85	-.63	.02
Mr. Stapleton	-.94	.10	.32
Dr. Grimesby Roylott	-.96	-.15	-.02

^a Decimals omitted.

high on the *Sophistication* dimension in company with the Napoleon of crime, Professor Moriarty. Finally, Conan Doyle is positioned midway between Holmes and his colleague, Dr. Watson. The character of Dr. Watson was apparently imbued with less estimable qualities than those ascribed to Conan Doyle by his admiring readers; that of Holmes, with more. Perhaps both characters are *personae* of Conan Doyle. As Baring-Gould (1968) observed: "... just as Conan Doyle in so many ways resembled Sherlock Holmes, so there are some remarkable parallels to be found in the lives of Conan Doyle and Dr. John H. Watson" (p. 10).

TABLE 4
CORRELATIONS BETWEEN SHERLOCK HOLMES DIMENSIONS AND RATING SCALES^a

	I Evaluative	II Sophistication	III Activity
Moral-Amoral	-.96	-.04	.05
Good-Evil	-.93	.14	.14
Cautious-Daring	-.48	-.62	.26
Weak-Strong	-.39	-.55	.38
Simple-Complex	-.38	-.81	-.05
Unintelligent-Intelligent	-.24	-.84	.29
Important-Unimportant	-.05	.82	.08
Active-Passive	.10	.24	-.67
Incompetent-Competent	.12	-.82	.01
Imaginative- Unimaginative	.29	.86	-.07
Flamboyant-Drab	.56	.41	.49
Sophisticated- Unsophisticated	.66	.55	.34
Ambitious-Unambitious	.70	.43	-.31

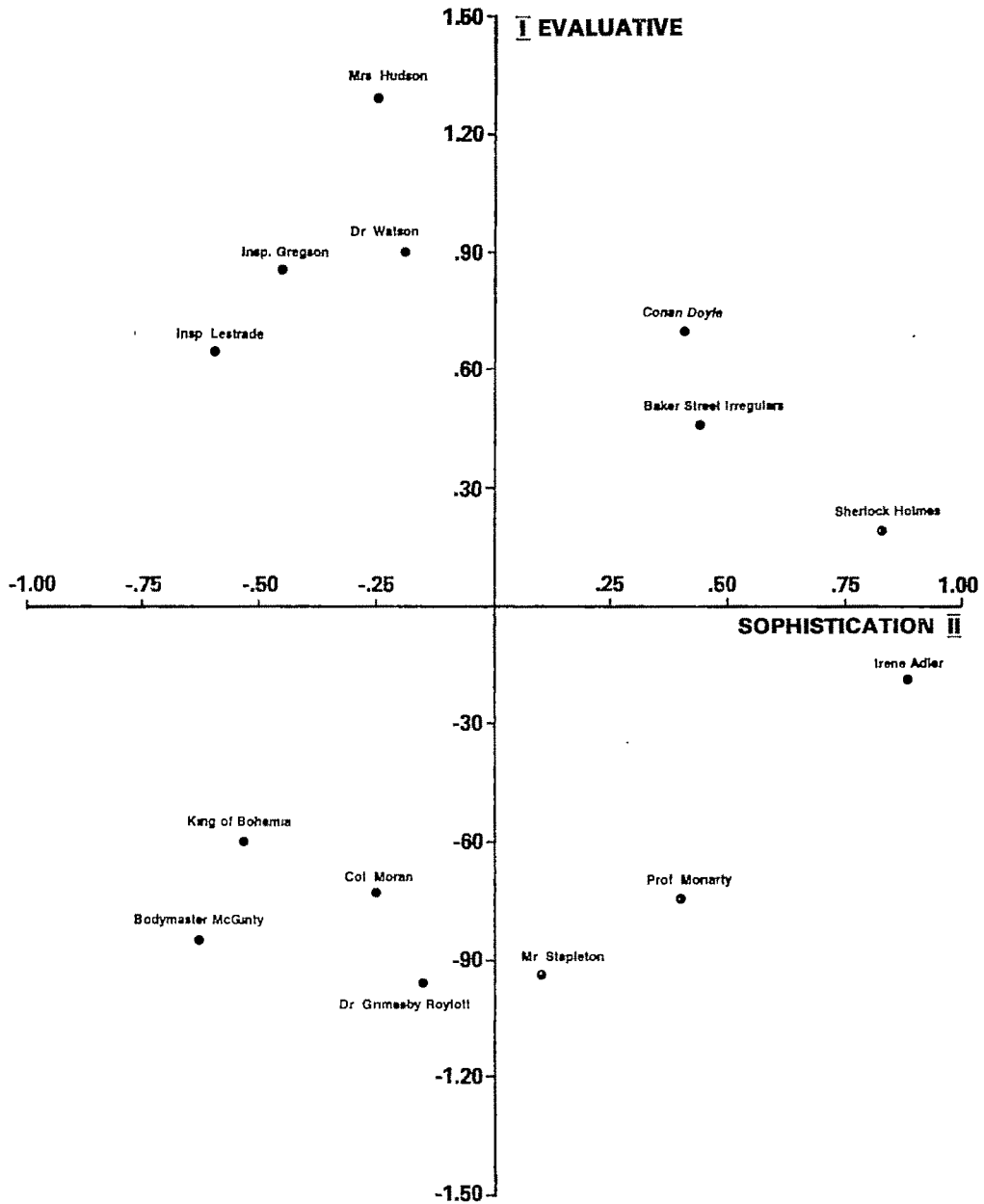
^a Decimals omitted.

DISCUSSION

In using characters as stimuli, we are probably dealing with changes in what was earlier classified by Wilson (1966) as interpretational responses to literature (i.e., the motivational forces and the nature of the characters). While his literary response category (i.e., plot mechanisms, language, style) undoubtedly also underwent changes, the nature and extent of change in this class of response was not assessed. For Othello, the initial factor loading on the Evaluative dimension was 0.09 whereas in the post-lecture analysis he has the highest positive loading (0.50) on the same dimension. Indeed, the *Othello* I and II Evaluative dimensions correlate 0.91 with the character of Othello included but correlate 0.97 with Othello removed from the comparison. Figure 1 shows the movement of Othello from a neutral position to that assigned the highest factor loading of any character. This fairly dramatic shift to a more positive regard for Othello is consistent with the emphasis in the lectures upon a sympathetic understanding of the universality of Othello's dilemma—a dilemma in which competing internal and external forces interlock to threaten annihilation.

Though one should caution against necessarily assuming an identity between dimensions given the same label, meaningful changes are evident in the Demeanour and Sophistication dimensions (Demeanour I and II correlated 0.71 while Sophistication I and II correlated 0.66). On both dimensions the sinister Iago has undergone a shift in spatial location being seen as more serious and a less sophisticated figure after detailed study. Moreover, he was relegated to an even more extreme position on the Evaluative dimension with the post-lecture judgments. Again, for example, the lectures drew attention to the self-destructive disintegration of Iago's plan and ultimately of Iago himself. Any lingering doubts as to his motives were dispelled and he emerges devoid of redeeming qualities. Figure 1 also shows Desdemona to have moved to a position of greater sophistication, probably because her speeches were seen to contain greater ironic foreshadowing in the later stages of the play. Particularly in the scene leading to her death she seems intuitively to understand the scheming of Iago and the part she has unwittingly played in the disintegration of Othello. The students were probably unaware of ironic subtleties until correlative ironies were detected in the speeches of the other major characters. Such nuances of irony would

FIGURE 2
CONAN DOYLE'S CHARACTERS IN TWO DIMENSIONAL SPACE



no doubt be overlooked in a first reading but would emerge sequentially. The lesser Montano and Gratiano underwent modest shifts in opposite directions for reasons that are not immediately apparent. Equally puzzling is the failure of Emilia to be seen as more sophisticated with study, particularly as she, like Desdemona, seems to utter speech laden with ironic subtleties which are more easily detected upon later consideration. Perhaps instructors treat the dynamics of major characters in disproportionately greater depth and leave the development of insight into the minor characters to the efforts of students—a forlorn hope.

While the above analyses speak only to how students perceive the characters within a work and do not by any means establish how the author intended they be seen, the suggestion is clear that with increased familiarity with a work different dimensions may be used (e.g., Control instead of Sexuality) and the order and distances among stimuli may change in ways consistent with lecture content. However, changes in the use of dimensions, particularly lesser dimensions, possibly reflect little more than instability in the solution and do not necessarily arise from the intervening teaching or study experience. At first blush the impact of the 'educational experience' seems convincingly demonstrated by the movement of stimuli (i.e., poems, characters) along the major dimensions in ways consistent with lecture content. However, an improved design for future studies would involve *predictions* of positional changes of characters in the configuration along stable dimensions established by this and other early investigations of the topic. In fact, that validation of educational impact was rather casually established in both the Berg and Wainer (1976) and the present study. True, certain key poems in the former study changed their dimensional positions according to lecture content as did the characters of Othello and Iago in this investigation. Whether others should have moved and did not, or, should not have and did, remains open to question.

The present design scarcely meets the criteria for a true experiment since it lacks a randomly comprised comparison group who were asked to make only the second set of dissimilarity judgments and adjective ratings. However, whether changes in perceptual structure arose from an artefact (e.g., pre-test sensitisation through administration of the initial dissimilarity scales) or as a result of the intervening lectures, discussion, and study treatment may not initially be of practical concern to the educator if the changes are such that the *combined* effect is to recast the characters in a different light or to bring out subtleties (i.e., intentions, weaknesses or strengths in individuals) among the relationships in directions considered 'correct' by the lecturer, relationships not immediately apparent to the students in their first reading of a work. However, if we are ultimately to relate changes in the order and distances of characters on dimensions, or, even to the use of different dimensions over the passage of time, specifically to a series of intervening lectures, a somewhat different design is called for. As a first step, equivalent groups could be asked to scale the same characters from a work before being exposed to lecturers differing radically in their interpretations and presentations of the piece. For example, a formalist presentation of Dreiser's *An American Tragedy* would tend to show the weakness of Clyde Griffiths as a character and the insufficiency of the resulting novel. However, a Marxist presentation would elevate Clyde as a noble character in conflict with a decadent capitalist society and praise the novel for its ideological purity. Similar changes in positions along an Evaluative dimension would be expected of secondary characters as for the protagonist, for example. This, or some similar procedure, would possibly establish the sensitivity of MDS techniques to educational experiences in a more convincing way. Validation by retrospective analysis should shortly give way to an empirical basis for assuming that the presentation of specific viewpoints would alter the reader's responses to the principal characters to some greater or lesser degree.

It should be noted that while the point of view emphasised in the lecture treatment

may effect changes in students' perceptual organisation, so too is the specific training and experience of the judges (students) themselves apt to have influenced their initial organisation. Distinct vantage points may be created though persistent influences brought to bear during the socialisation process, or indeed, through formal training such that information may be processed somewhat differently. For example, members of various denominations will organise a common domain of world religions somewhat differently (Russell *et al.*, 1979). Also, art professors tend to reflect the traditions of their schools in making evaluative judgments of student drawings (Klein, 1968). Thus, with some categories of literature (e.g., religious or political works) previously established personal or institutional points of view may be decisive in determining the structure of the reader's perceptual framework. Such a situation is probably the exception to the usual case where students are approaching individual works for the first time. One would assume the latter students would be more malleable and better able to profit from the 'educational' experience of the lecture hall insofar as they can acquire professorial perspectives without the need to first abandon less relevant schemas.

While it may be useful for an instructor to know the structure typically perceived to underlie an author's works or characters, the question of the generality of that structure across the creations of other writers is equally of interest. The Sherlock Holmes experts used as judges in exploring this question were unquestionably more knowledgeable about their subject matter than were the students at the *Othello II* stage. Discussion and detailed study over the years has developed a fine appreciation of many of the nuances and subtle themes woven throughout the stories. Though the subjects additionally differed in terms of age and culture (i.e., English vs. Canadian), and, though the two works are remarkably different in time, place, and genre, both shared a principal Evaluative dimension and a dimension labelled as Sophistication. Although the university students were asked to make judgments about characters from a single play, and the Sherlockians a sample of characters appearing in many stories, both dimensions dominated the respective MDS solutions. It is difficult to imagine two more disparate works, the one a play set in Renaissance Italy and the other a collection of Victorian novels and short stories. Yet, though one should be cautious in assuming an identity, Evaluative and Sophistication dimensions have been engaged by judges in three analyses in which the literary characters constituted the stimuli to be scaled. There seems to be no compelling reason for suggesting that we judge fictional characters any differently than we do those in real life. The dominance of an Evaluative dimension in a literary context merely supplements a wealth of existing data on person perception in which its emergence ahead of Potency and Activity dimensions usually results in factor analyses of the Semantic Differential scales (e.g., Warr and Knapper, 1968). However, with characters as stimuli, the fictive context created by a writer may, in rare instances, determine whether an Evaluative dimension can reasonably be expected to emerge. An author's unique portrayal of characters or the strong emphasis of a special theme may effect basic alterations in the student's perceptual organisation of the characters.

In the above, we have dealt with students' perceptual organisation of a work(s) and, perhaps indirectly, that of their professors through exposure to his point of view. A task for future research is to quantify an author's perception of his own work. Suffice it to say, it may or may not bear a resemblance to the structures of his readers or even to those of us (professors, actors, critics) who seek to interpret his work to the rest of the world. Individual difference models such as INDSCAL (e.g., Carroll, 1972) will allow an examination of individual points of view. The application of such procedures offers a valuable complement to the data base of future literary scholars in considering a writer's interpretations or comments about his own work.

ACKNOWLEDGMENTS.—The writers wish to thank members of The Sherlock Holmes Society of London who gave so generously of their time in assisting with this investigation. Specifically, the kindness of Lord Paul Gore-Booth, Philip Dalton, and Charles Scholefield, QC, is gratefully acknowledged.

REFERENCES

- BARING-GOULD, W. S. (Ed.) (1968). *The Annotated Sherlock Holmes*. Vol. 1. London: John Murray.
- BERG, W., and WAINER, H. (1976). Baudelaire in the classroom: an application of multidimensional scaling in educational evaluation. *Sci. Aesthetics*, 1, 1, 57-65.
- BERLYNE, D. E. (Ed.) (1974). *Studies in the New Experimental Aesthetics: Steps toward an Objective Psychology of Aesthetic Appreciation*. Washington, D.C.: Hemisphere.
- BERLYNE, D. E. (1976). The new experimental aesthetics and the problem of classifying works of art. *Sci. Aesthetics*, 1, 2, 85-106.
- CARROLL, J. D. (1972). Individual differences and multidimensional scaling. In SHEPARD, R. N., ROMNEY, A. K., and NERLOVE, S. G. (Eds.), *Multidimensional Scaling*. New York: Seminar Press.
- CRONBACH, L. J., and MEEHL, P. E. (1967). Construct validity in psychological tests. In JACKSON, D. N., and MESSICK, S. (Eds.), *Problems in Human Assessment*. New York: McGraw-Hill.
- FOREHAND, G. A. (1966). Problems of measuring response to literature. *The Clearing House*, 40, 6, 369-375.
- FUNK, S. G., HOROWITZ, A. D., LIPSHITZ, R., and YOUNG, F. W. (1976). The perceived structure of American ethnic groups: the use of multidimensional scaling in stereotype research. *Sociometry*, 39, 2, 116-130.
- HEILMAN, R. B. (1956). *Magic in the Web*. Lexington: University of Kentucky Press.
- JONES, L. E., and YOUNG, F. W. (1972). The structure of a social environment: a longitudinal individual differences scaling of an intact group. *J. pers. soc. Psychol.*, 24, 108-121.
- KLEIN, S. P. (1968). Using points of view and multidimensional scaling analyses to describe esthetic judgments. *Proceedings of the 76th Annual Convention of the American Psychological Association*, 3, 447-448.
- KRUSKAL, J. B., YOUNG, F. W., and SEERY, J. B. (1973). How to use KYST, a very flexible program to do multidimensional scaling and unfolding. Report No. 24, Bell Laboratories, Murray Hill, New Jersey.
- LINDAUER, M. S. (1974). *The Psychological Study of Literature*. Chicago: Nelson-Hall.
- MACCALLUM, R. C. (1974). Relations between factor analysis and multidimensional scaling. *Psychol. Bull.*, 81, 8, 505-516.
- ROMNEY, A. K., SHEPARD, R. N., and NERLOVE, S. B. (Eds.) (1972). *Multidimensional Scaling*. New York: Seminar Press.
- ROSENBERG, S., and JONES, R. (1972). A method for investigating and representing a person's implicit theory of personality: Theodore Dreiser's view of people. *J. Pers. soc. Psychol.*, 22, 372-386.
- RUSSELL, G. W., GODDARD, A. M., and PENTON, M. J. (1979). The perception of Judeo-Christian religions. *Canad. J. behav. Sci.*, 11(2), 140-152.
- WAGENAAR, W. A., and PADMOS, P. (1971). Quantitative interpretation of stress in Kruskal's multidimensional scaling technique. *Br. J. Math. statist. Psychol.*, 24, 101-110.
- WAINER, H., and BERG, W. (1972). The dimensions of de Maupassant: a multidimensional analysis of students' perception of literature. *Am. educ. res. J.*, 9, 4, 485-491.
- WAINER, H., and KAYE, K. (1974). Multidimensional scaling of concept learning in an introductory course. *J. educ. Psychol.*, 66, 4, 591-598.
- WARR, P. B., and KNAPPER, C. (1968). *The Perception of People and Events*. London: Wiley.
- WILSON, J. R. (1966). Responses of college freshmen to three novels. Report No. 7, National Council of Teachers of English, Champaign, Illinois.
- WINER, B. J. (1971). *Statistical Principles in Experimental Design*. New York: McGraw-Hill.
- YOUNG, F. W. (1968). A FORTRAN IV program for nonmetric multidimensional scaling. Report No. 56, L. L. Thurstone Psychometric Laboratory, University of North Carolina, Chapel Hill, North Carolina.

(Manuscript received 2nd November, 1979)

EFFECTS OF PERSONAL LECTURE NOTES AND TEACHER-NOTES ON RECALL OF UNIVERSITY STUDENTS

By M. MAQSUD

(Faculty of Education, Bayero University, Kano, Nigeria)

SUMMARY. Two experiments studied the effects of note-taking, strategy of note-taking (short/long notes) and reviewing personal notes and/or simplified teacher-notes on immediate and delayed recall. One hundred and sixty university students, classified as either short or long note-takers, served as subjects. They listened to a recorded lecture and then their immediate and delayed retention was assessed by means of the free recall test. Analysis of results revealed that the act of note-taking has facilitative effects on both immediate and long-term recall. Free recall scores of subjects who took short notes in their normal lectures tended to be significantly higher than those who took long notes. Reviewing personal lecture notes and/or simplified and organised teacher-notes facilitates delayed recall.

INTRODUCTION

It is believed that the act of note-taking usually helps university students retain lecture material. The effects of note-taking on memory are generally attributed to an encoding function (Howe, 1970) or an external storage function (Anderson, 1970). Howe (1970) presented the encoding hypothesis to explain facilitative effects of note-taking on recall. The hypothesis suggests that note-taking involves encoding function by coding, naming, chaining, and categorising material in terms of some pre-established scheme (Howe, 1976); students transform learning material into subjective and meaningful forms (Howe, 1977). Anderson (1970) emphasised that external storage is the sole factor aiding recall; lecture notes may function as an external storage to preserve information for later review.

Some investigators (Di Vesta and Gray, 1972; Fisher and Harris, 1973) showed positive effects of note-taking on recall, while others (Berliner, 1969; Peters, 1972) did not see facilitative effects of note-taking on retention. Carter and Van Matre (1975) pointed out that the conflicting results were due to differences in three variables: opportunities for reviewing notes, timing of review before the test, and duration of the retention interval. In their experimental study, Carter and Van Matre found that the condition of taking notes while listening to the recorded lecture and reviewing notes immediately before the free recall test produced maximum retention and recall efficiency. Supporting the external storage function, Carter and Van Matre attributed facilitative effects of note-taking to having opportunities for reviewing notes, and concluded that 'note-taking when coupled with notes review represents a useful strategy for affecting subsequent recall; note-taking alone is of questionable value' (Carter and Van Matre, 1975, p. 903).

Learners employ a variety of procedures and strategies to retain learning material; they use a number of ways to make material meaningful to them (Howe, 1976). It is a matter of common observation that university students differ in employing strategies of note-taking. Some students take detailed notes and others take brief notes while listening to lectures. The former tend to make their external storage as large as they can, with their intention to keep external storage for future use; they integrate new material into the existing one when they review their notes, usually before taking examinations. And the latter tend to make new material meaningful to them while they take notes. Both strategies of note-taking in fact imply the process of encoding lecture material, but at different times.

Following the encoding hypothesis, students who encode material while listening to lectures are expected to show superior performance on an immediate free recall test. University students, in addition to their own personal lecture notes, ask for teacher-notes, simplified and organised notes prepared by their lecturers. Before taking examinations, students review their own notes and teacher-notes so that they can conveniently recall factual material. Since retention plays an important role in the existing examination systems, it seems desirable to see whether review of personal notes and/or teacher-notes has some significant effects on delayed recall.

METHOD

EXPERIMENT I

Subjects and design

Eighty male undergraduate students of Educational Psychology, studying at a Nigerian university, took part in the experiment. Their personal lecture notes for three one-hour lectures in Educational Psychology were collected as they left lecture rooms and then carefully examined by the researcher and his two assistants. On the basis of the number of words written down in their personal notes for three lectures, subjects were classified as brief note-takers (average mean = 445 words per one-hour lecture) and detailed note-takers (average mean = 744 words per one-hour lecture), 40 subjects under each group, split at the median. According to their scores on the Mill Hill Vocabulary Scale (Raven, 1965), each of the two main groups was divided into two subgroups, high (Hi) and low (Lo) split at the medians (see Table 1). Forty subjects, 10 chosen randomly from each of the four subgroups, were asked to take notes, whereas 40 other subjects did not take notes while listening to a lecture. There were three independent variables (note-taking, strategy of note-taking and verbal ability) and a dependent variable (free recall); all these variables entered into a $2 \times 2 \times 2$ factorial design.

TABLE 1
MEANS AND STANDARD DEVIATIONS OF VOCABULARY SCORES

Group	Strategy of note-taking (Total Sample = 80)			
	Short (Total N = 40)		Long (Total N = 40)	
	Mean	SD	Mean	SD
Hi	52.20	3.99	47.30	4.82
Lo	38.00	4.72	33.75	3.60

Materials

A lecture, consisting of 2,241 words, on 'Emotional Development' was written by the experimenter. Two other lecturers in Educational Psychology considered the lecture material suitable and corresponding with the level of understanding of subjects. Since subjects were quite familiar with the experimenter's voice, the lecture was recorded in his voice and presented at about 110 words per minute. To assess the verbal ability, the Mill Hill Vocabulary Scale was used. For the free recall test, subjects were asked to write as much as they could recollect about the recorded lecture.

Procedure

To take the measures of verbal ability, all subjects were seated in a lecture hall and asked to complete the Mill Hill Vocabulary Scale. A week after the administration of the vocabulary scale, 40 subjects, selected for note-taking treatment, were again

seated in a lecture hall. They were provided with sheets of paper and pencils and instructed to take notes while listening to the recorded lecture, as they did during normal lectures. These subjects were, however, not informed of taking a free recall test immediately after the lecture. As soon as the recorded lecture ended, subjects' notes were collected by the experimenter and they were given some more sheets of paper for the free recall test. Subjects were then told to write as much as they could remember about the lecture. No time limit was imposed and subjects were allowed to leave the testing room whenever they felt that they had completed the free recall test.

The second group of 40 subjects were also seated in the lecture hall. They listened to the recorded lecture but were told not to take notes and they were also unaware of taking the free recall test. Immediately after the lecture, these subjects, too, were asked to write about the lecture material.

RESULTS (EXPERIMENT I)

A key, listing 741 identifiable information units of the lecture material, was prepared by the experimenter and two lecturers in Educational Psychology. For scoring purposes, each identifiable information unit scored 1. The experimenter and his two assistants scored subjects' free recall material independently. Inter-rater reliability ranged between 0.89 and 0.95. Mean free recall scores together with SDs are given in Table 2.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF FREE RECALL SCORES
(IMMEDIATE RECALL)

Note-taking	Vocabulary	Strategy of note-taking (Total Sample = 80)			
		Short (Total N = 40)		Long (Total N = 40)	
		Mean	SD	Mean	SD
Yes	Hi	68.70	5.14	57.20	7.45
	Lo	56.10	5.26	51.20	4.69
No	Hi	48.30	3.26	40.80	5.02
	Lo	45.80	5.51	37.30	3.29

The data were analysed by means of a three-way analysis of variance. The analysis yielded significant main effects for note-taking, $F(1, 72) = 181.25$, $P < 0.001$, strategy of note-taking, $F(1, 72) = 51.13$, $P < 0.001$, and verbal ability, $F(1, 72) = 29.48$, $P < 0.001$. It can be interpreted that note-taking and high verbal ability have facilitative effects on immediate recall. The strategy of taking short notes is more effective than that of taking long notes.

Among the first order interactions, only verbal ability and note-taking reached the conventional level of significance, $F(1, 72) = 7.73$, $P < 0.01$. The second order interaction (note-taking $\times$ strategy of note-taking $\times$ verbal ability) also showed significant effects on immediate recall, $F(1, 72) = 12.08$, $P < 0.001$. These results revealed that immediate recall of subjects who had high verbal ability and used the strategy of taking short notes was superior to that of those who had low verbal ability and took long notes while listening to their lectures.

EXPERIMENT II

Results of Experiment I revealed that note-taking has facilitative effects on immediate recall. Those subjects who took short notes during their normal lectures

tended to show superior immediate recall to that of those who took long notes. Generally, these results were in support of the encoding hypothesis—short note-takers made the lecture material meaningful to them. The aims of Experiment II were to replicate some of the findings of Experiment I, and also to see the effects of reviewing notes and having an access to relevant material (teacher's notes) on delayed recall.

Subjects and design

For this experiment, another group of 80 male undergraduate students, registered for the course of Developmental Psychology, served as subjects. Only two factors, strategy of note-taking and different conditions of reviewing the lecture material immediately before the free recall test, were studied and these between-subject factors entered into a 2×4 analysis of variance.

Materials

The recorded lecture used for Experiment I was also used for Experiment II, as subjects of this study did not take part in Experiment I. The experimenter prepared a detailed but simplified and organised hand-out. Two judges, lecturers in Educational Psychology, read the experimenter's hand-out and made certain that no additional information unit was included in the hand-out.

Procedure

The experimenter and his two assistants examined subjects' lecture notes for their three normal lectures they attended for the course of Developmental Psychology. On the basis of the number of words written down in their personal notes for three one-hour lectures, subjects were classified as short note-takers (average mean = 461 words per one-hour lecture) or long note-takers (average mean = 783 words per one-hour lecture). Following the design of the experiment, subjects under each of the two categories, split at the median (40 short note-takers and 40 long note-takers), were randomly distributed into four sub-groups, 10 subjects in each group. Four groups (40 subjects), two groups from each of the two categories (short/long note-takers), were seated in a lecture hall and they were told to take notes, while listening to the recorded lecture. When the lecture ended, these subjects were told that they would be required to take a test after seven days. Four other groups were also seated in a lecture hall, but they were instructed not to take notes while listening to the recorded lecture. These subjects were also asked to report for the test after seven days.

To study the effects of different conditions of reviewing before free recall test on retention, each subgroup had to follow one of the four conditions of reviewing: personal lecture notes, personal lecture notes plus teacher's hand-out, teacher's hand-out, and mental review. Subjects were seated in a hall and asked to do reviewing for 30 minutes. Three hours after the period of review, subjects were again seated in the hall and asked to write as much as they could recollect about the content of the recorded lecture. The procedure of scoring was the same as reported for Experiment I.

RESULTS (EXPERIMENT II)

Table 3 gives mean free recall scores and SDs for four conditions of reviewing before taking the free recall test. Free recall scores were submitted to a two-way analysis of variance. The main effects for strategy of note-taking (short/long notes), $F(1, 72) = 4.22$, $P < 0.05$, and condition of reviewing, $F(3, 72) = 459.32$, $P < 0.001$, were found significant. But the interaction between the strategy of note-taking and condition of reviewing, $F(3, 72) = 1.96$, NS, was not found significant. It could be seen from Table 3 that effects of the four conditions of reviewing on free recall scores were consistent for both short and long note-takers. Among the four conditions of

TABLE 3
MEANS AND STANDARD DEVIATIONS OF FREE RECALL SCORES (DELAYED
RECALL)

Note-taking	Condition of reviewing	Strategy of note-taking (Total Sample = 80)			
		Short (Total N = 40)		Long (Total N = 40)	
		Mean	SD	Mean	SD
Yes	Personal notes	39.20	2.22	39.10	2.12
Yes	Personal notes plus teacher's hand-out	44.80	2.79	42.00	2.45
No	Teacher's hand-out	42.20	1.27	41.80	1.25
No	Mental	22.70	1.27	22.40	1.11

reviewing, 'personal notes plus teacher's hand-out' was the most effective, followed by 'teacher's hand-out' and then 'personal notes'; 'mental reviewing' was the least effective. As far as strategy of note-taking is concerned, only the condition of 'personal notes plus teacher's hand-out' yielded significant variations between short and long note-takers ($t = 2.39$, $df = 18$, $P < 0.05$).

DISCUSSION

The results of Experiment I showed that the act of taking notes, without any review, has facilitative effects on immediate recall. But mean free recall scores for subjects, who were classified as brief note-takers on the basis of their notes for three normal lectures, were significantly higher than those for subjects who were classified as detailed note-takers. These results tend to support the encoding hypothesis that note-taking implies transformation of learning materials into meaningful structures. If notes are meant exclusively for external storage, the act of taking notes did not have significant effects on immediate free recall scores. The results of Experiment I are in line with the findings of Di Vesta and Gray (1972) and that of Fisher and Harris (1973). Our interpretation, however, contradicts the findings of Berliner (1969) and Peters (1972); no positive effects of note-taking on retention were seen by these researchers.

Carter and Van Matre (1975) concluded that notes function as an external storage. As far as the effect of note-taking on immediate recall is concerned, the design of Carter and Van Matre's study does not seem sound. By placing the review session before the free recall test, the results, in favour of the external storage function, were expected. To examine the effects of encoding function on immediate recall, the free recall test should in fact be given to subjects immediately after they finish listening to lecture material and note-taking without any review session.

A positive association between high verbal ability and free recall scores was expected. Vocabulary is an internal verbal storage of a person. Verbal learning material can conveniently be linked with the existing knowledge and made meaningful through verbal mediation if the learner has a broader verbal vocabulary.

Apparently, the results of Experiment II tend to support the external storage function (Anderson, 1970) if we regard the teacher's hand-out as an external source of information. But we believe that learners tend to encode new learning material (Howe, 1976). The teacher's hand-out might have given some additional points, not taken earlier, to subjects who read the hand-out before taking the free recall test, and thus raising their free recall scores.

The superiority of short note-takers who reviewed their personal notes plus teacher's hand-out tends to substantiate the findings of Experiment I. The results of Experiment II suggest that short note-takers tend to transform learning material into subjective forms more rapidly and they not only apply this habitual approach to auditory verbal learning material but also to reading material. The review of the teacher's hand-out perhaps gave them some additional cues which were quickly linked with the structures they had already made.

The results of Experiment II emphasise the point that teacher-notes facilitate recall of university students in examination-like situations, provided these notes are presented in simplified and well-organised form (Tulving and Donaldson, 1972). These findings are supported by general observations that university students attach much importance to teacher-notes and often make demands for such notes.

It is concluded that university students tend to encode new learning material while listening to lectures and there are different strategies of encoding; some strategies are more facilitative than others to both immediate and delayed recall. Organised and simplified teacher-notes help students in examination-like situations where retention plays a significant role.

REFERENCES

- ANDERSON, R. C. (1970). Control of student mediating process during verbal learning and instruction. *Rev. educ. Res.*, 40, 349-369.
- BERLINER, D. C. (1969). Effects of test-like events and note taking on learning from lecture instruction. *Proc. 77th Ann. Convent. Am. Psychol. Assoc.*, 4, 623-624.
- CARTER, J. F., and VAN MATRE, N. H. (1975). Note-taking versus note-taking. *J. educ. Psychol.*, 67, 900-904.
- DI VESTA, F. J., and GRAY, S. G. (1972). Listening and note-taking. *J. educ. Psychol.*, 63, 8-14.
- FISHER, J. L., and HARRIS, M. B. (1973). Effects of note-taking and review on recall. *J. educ. Psychol.*, 65, 321-325.
- HOWE, M. J. A. (1970). Using student's notes to examine the role of the individual learner in acquiring meaningful subject matter. *J. educ. Res.*, 64, 61-63.
- HOWE, M. J. A. (1976). Good learners and poor learners. *Bull. Br. Psychol. Soc.*, 29, 16-19.
- HOWE, M. J. A. (1977). *Adult Learning: Psychological Research and Applications*. London: Wiley.
- PETERS, D. L. (1972). Effects of note-taking and rate of presentation on short-term objective test performance. *J. educ. Psychol.*, 63, 276-280.
- RAVEN, J. C. (1965). *Guide to Using the Mill Hill Vocabulary Scale*. London: H. K. Lewis.
- TULVING, E., and DONALDSON, W. (1972). *Organization of Memory*. New York: Academic Press.

(Manuscript received 21st January, 1980)

RESEARCH NOTES

THE USE OF ORIENTATIONS IN CHILDREN'S DRAWINGS OF FAMILIAR OBJECTS: PRINCIPLES VERSUS PERCEPTS

By S. W. IVES

(*Wheelock College, Boston, and Harvard Project Zero*)

SUMMARY. It has been suggested that drawings are constructions which follow certain principles and that young children are reluctant to relinquish these principles. Kindergartners, second graders, fourth graders were asked to draw familiar objects (e.g., toy car) placed in several orientations or in an orientation of their choice to determine if the orientation depicted in their drawings would be guided more by graphic principles or by the view before them. The use of graphic principles prevailed in the choice of orientation even though many children in second and fourth grade carefully recorded other aspects particular to the objects in front of them (e.g., make and model of car).

INTRODUCTION

Arnheim (1974) suggests that drawings are not simply copies of either perception or knowledge but rather equivalent constructions within the bounds of the medium. Others have gone on to search for construction rules in children's drawings (e.g., Willats, 1977); one initial finding is that prior to third grade, children seem to have difficulty relinquishing such graphic principles as the arrangement of parts on a horizontal-vertical axis (Goodnow, 1978) or the consistent use of standard orientations for certain objects, e.g., front view for a person, side view for a horse, even when the task calls for a change in orientation (Ives and Houseworth, 1980; Ives and Rovet, 1979). These latter studies, however, asked children to draw without an actual model present. Freeman and Janikoun (1972) found that children prior to age seven will draw a cup which only provides the essential defining features regardless of the nature of the cup actually placed before them. The present study seeks to determine whether children will be reluctant to relinquish the use of standard orientations when depicting familiar objects placed before them.

METHOD

Sample

Thirty-six children were involved (equal numbers of boys and girls and kindergartners, second, and fourth graders; mean ages 5·8 years, 7·5 years, and 9·8 years). All children attended an elementary school in a low-income neighbourhood of a major city in Ontario, Canada. Testing was carried out in a single 30-minute session for each child.

Apparatus and procedure

Four stimulus objects were used: a toy car, 5 cm high by 18 cm long; a toy man, 18 cm high by 5 cm wide; a toy house, 14 cm high by 20 cm long; a toy horse, 16 cm high by 18 cm long. Each toy was realistic in design. Each child had to draw all four objects in one of three conditions: the toy was placed by the experimenter in either a front or a side orientation in relation to the child or placed by the children in an orientation of their choice.

RESULTS AND DISCUSSION

Objects were scored as being drawn in either a front, side, or mixed (front and side) orientation by two judges. Interjudge reliability was better than 95 per cent with differences settled by consensus. A 3 (grade level) $\times$ 3 (placement condition) $\times$ 4 (objects) analysis of variance was carried out on the number of front view drawings. The object effect is significant, $F_{(3, 12)} = 83.36$, $P < 0.001$ (83 per cent of the total variance using Dwyer's (1974) method for testing the magnitude of effects), with the man and house primarily drawn in the front view (67 per cent) and the horse and car in the side view (83 per cent) regardless of the orientation in which they are placed before the child. There is also a significant grade level effect, $F_{(2, 12)} = 10.84$, $P < 0.005$ (5 per cent of the variance) and a significant grade level $\times$ object interaction, $F_{(6, 12)} = 3.61$, $P < 0.05$ (5 per cent of the variance) primarily reflecting the fact that the house tends to be drawn in the front view by kindergartners (75 per cent) while the older children tend to depict the house in a combination of the front and side view (63 per

cent) which was scored as mixed. The placement condition had no significant effect on the orientation drawn. In addition, the object had a significant effect on the child's own placement, $F_{(3,6)} = 24.00$, $P < 0.001$ (92 per cent of the variance) with free-choice placements following the overall drawing patterns, man and house in the front view (63 per cent) and car and horse in the side view (83 per cent).

The drawing and placement orientations in the present study are consistent with previous findings (Ives and Houseworth, 1980; Ives and Rovet, 1979). However, this does not necessarily mean that children ignore all perceptual information simply to draw what they know of the basic category 'car', for example, as Piaget and Inhelder (1956) suggest. While only 33 per cent of the kindergartners carefully examined the object as they drew it, all of the second and fourth graders did so and these older children reproduced many details (e.g., make and model of the car) yet changed the orientation to fit the standard form. Young children will thus construct drawings in a manner consistent with graphic principles or rules even when confronted with contrary information before them.

ACKNOWLEDGMENTS.—This research was submitted in partial fulfilment of the degree of Ph.D at the University of Toronto. The author would like to thank Drs. D. R. Olson and J. Rovet for advice and encouragement.

REFERENCES

- ARNHEIM, R. (1974). *Art and Visual Perception: The New Version*. Berkeley: University of California Press.
- DWYER, J. H. (1974). Analysis of variance and the magnitude of effects: A general approach. *Psychol. Bull.*, 81, 731-737.
- FREEMAN, N. H., and JANIKOUN, R. (1972). Intellectual realism in children's drawings of a familiar object with distinctive features. *Child Dev.*, 43, 1116-1121.
- GOODNOW, J. J. (1978). Visual thinking: cognitive aspects of change in drawings. *Child Dev.*, 49, 637-641.
- IVES, W. W., and HOUSEWORTH, M. (1980). The role of standard orientations in children's drawings of interpersonal relationships: aspects of graphic feature marking. *Child Dev.*, 51, 591-593.
- IVES, S. W., and ROVET, J. (1979). The role of graphic orientations in children's drawings of familiar and novel objects at rest and in motion. *Merrill-Palmer Quart.*, 25, 281-292.
- PIAGET, J., and INHELDER, B. (1956). *The Child's Conception of Space*. London: Routledge and Kegan Paul.
- WILLATS, J. (1977). How children learn to draw realistic pictures. *Quart. J. Exper. Psychol.*, 29, 367-382.

(Manuscript received 10th March, 1980)

CHURCH AIDED SECONDARY SCHOOLS, RELIGIOUS EDUCATION
AS AN EXAMINATION SUBJECT AND PUPIL ATTITUDE
TOWARDS RELIGION

By L. FRANCIS
(*London Central YMCA*)

AND
MARIAN CARTER
(*Wentworth Castle College of Education, Barnsley*)

SUMMARY. Attitudes towards religion of 802 fifth-year secondary pupils were studied in relation to whether they attended a church voluntary aided or Local Education Authority school and whether they belonged to a CSE/GCE examination group in religious education. Comparisons were made between pupils in Local Education Authority schools taking RE as an examination subject, pupils in Local Education Authority schools not taking RE, pupils in Church of England voluntary aided schools taking RE, pupils of Church of England voluntary aided schools not taking RE and pupils in Roman Catholic voluntary aided schools taking RE. It was found that there was no difference in pupil attitude in the Church of England RE and non-RE groups, the Roman Catholic group and the Local Education Authority non-RE group. The Local Education Authority RE group scored significantly higher on the attitude scale. These findings provide no support for the notion that church voluntary aided secondary schools exert a positive influence on their pupils' attitudes towards religion.

INTRODUCTION

As a consequence of the 1944 Education Act, the churches in England have the option of providing voluntary aided schools which are able to select pupils on religious criteria and in which denominational religious education can be given (Dent, 1947; Alexander and Taylor, 1977). The two churches which have placed considerable investment in the voluntary aided system are the Roman Catholic Church and the Church of England. Both the Roman Catholic Church (Jebb, 1968) and the Church of England (Durham Report, 1970) have discussed the significance of voluntary aided schools in terms of the formation of pupil attitudes. It is reasonable to hypothesise that if the church voluntary aided schools are achieving what they set out to achieve, pupils in church voluntary aided schools should possess a more favourable attitude towards religion than pupils in state schools.

Francis (1979) reports on the influence of Roman Catholic and Church of England voluntary aided primary schools on the attitude of third- and fourth-year pupils towards religion. School influence is discussed only after a number of other potentially contaminating variables are taken into account through a sequence of developing path models, namely sex, age, social class, IQ, religious behaviour and parental religious behaviour. Francis demonstrates that pupils educated in Roman Catholic voluntary aided primary schools have a more favourable attitude towards religion than children in Local Education Authority schools. However, pupils educated in Church of England voluntary aided primary schools do not have a more favourable attitude towards religion than children in LEA schools. The present paper employs the same attitude scale among fifth-year secondary pupils in LEA, Roman Catholic voluntary aided and Church of England voluntary aided secondary schools to examine whether church voluntary aided secondary schools have the same relationship to pupil attitude as church voluntary aided primary schools.

At the fifth-year secondary level the situation is complicated by the fact that some pupils take religious education as a school examination subject and devote to it on average four periods per week, while other pupils do not. In the LEA schools in the sample the majority of pupils did not take religious education as an examination subject. In the Church of England voluntary aided schools only a minority of pupils did not take religious education as an examination subject. In the Roman Catholic voluntary aided schools all the pupils took religious education as an examination subject. Consequently, a comparison can be made between pupils educated according to five different environments of religious education:

(1) children taking religious education as an examination subject in Local Education Authority schools;

(2) children not taking religious education as an examination subject in Local Education Authority Schools;

(3) children taking religious education as an examination subject in Church of England voluntary aided schools;

(4) children not taking religious education as an examination subject in Church of England voluntary aided schools;

(5) children taking religious education as an examination subject in Roman Catholic voluntary aided schools.

Following Francis (1979), a number of variables were also taken into consideration because of their known correlation with pupil attitudes towards religion: sex (Garrity, 1960; Povall, 1971), intelligence (Hyde, 1959), religious behaviour (Johnson, 1966), parental religious behaviour (Greer, 1971), socio-economic background (Jones, 1962), extraversion and neuroticism (Carter, 1979).

METHOD

Sample

Twenty-one secondary schools were selected from the three areas of Essex, Yorkshire and Nottinghamshire. In each of the schools the questionnaire was administered to one fifth-year class which was taking religious education as an examination subject. In the LEA and Church of England voluntary aided schools the questionnaire was also administered to a comparable class which was not taking religious education as an examination subject. Eight hundred and two fifth-year secondary pupils took part in the survey.

Measures

(1) Scale of attitude towards religion Form ASC4B (see Francis, 1978, 1979). The scale contains items concerned with the child's attitude towards God, Jesus, the Bible, prayer, the church and religion in schools. The scale is known to function reliably (alpha coefficient = 0.97) and validly among fifth-year secondary pupils.

(2) Frequency of pupil's church attendance, pupil's prayer, mother's church attendance and father's church attendance. These variables are gauged on a four-point scale ranging from weekly to never.

(3) Socio-economic groups. The scale proposed by the Office of Population Censuses and Surveys (1970) is used in association with information about parental employment.

(4) IQ rating. IQ is measured by Raven's Progressive Matrices (Raven, 1947).

(5) Extraversion and neuroticism. These are measured by the Junior Eysenck Personality Inventory (Eysenck, 1965).

Procedure and data analyses

The questionnaires were all administered by the same researcher who followed a standardised procedure and the data analysed by means of linear multiple regression and path analysis using the SPSS package (Nie *et al.*, 1975).

RESULTS

The inter-relationships between the predictor variables were explored through a developing sequence of path models analogous to the sequence proposed by Francis (1979). The present brief report presents only the final stages of this sequence as sufficient to demonstrate the relationship between the school type and pupil attitude towards religion.

Table 1 examines the increased proportion of variance in the scores of attitude towards religion which is accounted for when the school type is entered into the multiple regression equation. This variable is entered into the equation through the creation of a set of dummy variables (Cohen, 1968) in which the Local Education Authority non-RE group functions as the reference category. School type is significant at the 0.025 level of probability.

Table 2 examines the significance of the separate contributions of each of the four school types to the prediction of scores of attitude towards religion. From this table it appears that only the Local Education Authority RE group contributes significant predictive power to the equation. The regression coefficient indicates that the predicted attitude scores for children

TABLE 1
BLOCKWISE MULTIPLE REGRESSION SIGNIFICANCE TESTS

Predictors	Total R^2	Increase in R^2	F Ratio	df	P <
Block one	0.57785	0.57785	109.213	9,697	0.001
Block two	0.59024	0.01239	5.269	4,697	0.025

- Notes: 1. Block one contains the predictors: sex, social class, extra-version, IQ, neuroticism, mother's church attendance, father's church attendance, frequency of praying and frequency of church attendance.
2. Block two contains four dummy variables representing the Local Education Authority RE group, the Church of England non-RE group, the Church of England RE group and the Roman Catholic RE group. The Local Education Authority non-RE group serves as the reference category.

TABLE 2
MULTIPLE REGRESSION SIGNIFICANCE TESTS

Predictor	Total R^2	Increase in R^2	B	Beta	St. error B	F ratio	df	P <
Block one	0.57785	—	—	—	—	—	—	—
LEA, RE group	0.58609	0.00824	+5.07625	+0.11137	1.34724	14.197	1,697	0.001
C of E, non-RE group	0.58835	0.00226	-2.98361	-0.03520	2.19795	1.843	1,697	NS
C of E, RE group	0.58900	0.00065	+2.74248	+0.04289	1.78812	2.352	1,697	NS
Roman Catholic group	0.59024	0.00124	+2.24779	+0.04430	1.54800	2.108	1,697	NS
CONSTANT			118.6217					

Note. The Local Education Authority non-RE group serves as the reference category.

in the Local Education Authority RE group would be about five points higher on the 96-point attitude scale than for comparable children in the Local Education Authority non-RE group.

CONCLUSION

In the case of fifth-year secondary pupils, neither Roman Catholic nor Church of England voluntary aided secondary schools seem to be achieving a positive influence on their pupils' attitude towards religion. Whereas Francis (1979) suggested that Roman Catholic voluntary aided primary schools were exerting a positive influence on the attitudes of third- and fourth-year pupils towards religion, the present study indicates that this positive influence is no longer the case among fifth-year secondary pupils. The Roman Catholic church seems to have been more successful in influencing pupil attitude towards religion through its primary voluntary aided schools than through its secondary voluntary aided schools. On the other hand, the Church of England seems to have no positive influence on its pupils' attitude towards religion either through its primary or secondary voluntary aided schools.

One further feature of the data requires comment, namely the significant positive relationship between the provision of religious education as an examination subject in LEA schools and favourable attitudes towards religion, and the lack of such relationship in the church voluntary aided schools. In the church voluntary aided schools in the sample religious education was taken as an examination subject by the majority of pupils, whether they would have wished to opt for the subject or not. In the LEA schools religious education as an examination subject was clearly a minority option. It is reasonable to assume that the pupils who took the subject option in the Local Education Authority schools had a clear interest in it. In this situation the provision of the subject option seems to interact with pupil disposition to produce a minority group of pupils who report significantly more favourable attitudes towards religion. In the church voluntary aided schools, on the other hand, there is

no similar opportunity for distinguishing and teaching separately a group which has a special interest in religion. If one of the objectives behind the provision of compulsory (or semi-compulsory) religious education as an examination subject in church voluntary aided secondary schools is the development in their pupils of a more favourable attitude towards religion than that held by pupils in Local Education Authority schools who do not take religious education as an examination subject, this objective seems to have failed.

REFERENCES

- ALEXANDER, W., and TAYLOR, G. (1977). *County and Voluntary Schools*. (5th ed.) London: Councils and Education Press.
- CARTER, M. (1979). The development of aspects of self in adolescence with reference to religious and moral education. Unpublished M.Phil. dissertation, University of Nottingham.
- COHEN, J. (1968). Multiple Regression as a general data analytic system. *Psychol. Bull.*, 70, 426-443.
- DENT, H. J. (1947). *The Education Act 1944: Provisions, Possibilities and Some Problems*. (3rd ed.). London: University of London Press.
- DURHAM REPORT. (1970). *The Fourth R*. London: SPCK and NS.
- EYSENCK, S. B. G. (1965). *Manual of the Junior Eysenck Personality Inventory*. London: University of London Press.
- FRANCIS, L. J. (1978). Attitude and longitude: a study in measurement. *Character Potential*, 8, 119-130.
- FRANCIS, L. J. (1979). School influence and pupil attitude towards religion. *Br. J. educ. Psychol.*, 49, 107-123.
- GARRITY, F. D. (1960). A study of the attitude of some secondary modern school pupils towards religious education. Unpublished M.Ed. dissertation, University of Manchester.
- GREER, J. E. (1971). Religious belief and church attendance of sixth form pupils and their parents. *Irish J. Educ.*, 5, 98-106.
- HYDE, K. E. (1959). A study of some factors influencing the communication of religious ideas and attitudes among secondary schoolchildren. Unpublished Ph.D. dissertation, University of Birmingham.
- JEBB, P. (Ed.) (1968). *Religious Education: Drift or Decision?* London: Darton, Longman and Todd.
- JOHNSON, W. P. C. (1966). The religious attitude of secondary modern county school pupils. Unpublished M.Ed. dissertation, University of Manchester.
- JONES, J. A. (1962). An investigation into the response of boys and girls to scripture as a school subject in certain co-education grammar schools in industrial South Wales. Unpublished M.A. dissertation, University of Wales (Swansea).
- NIE, N. H., HULL, C. H., JENKINS, J. G., STEINBRENNER, K., and BENT, D. H. (1975). *Statistical Package for the Social Sciences*. (2nd ed.) New York: McGraw-Hill.
- OFFICE OF POPULATION CENSUSES AND SURVEYS (1970). *Classification of Occupations*. London: HMSO.
- POVALL, C. H. (1971). Some factors affecting pupils' attitudes to religious education. Unpublished M.Ed. dissertation, University of Manchester.
- RAVEN, J. C. (1947). *Progressive Matrices*. London: H. K. Lewis.

(Manuscript received 26th October, 1979)

A SIMPLE FORM OF TAILORED TESTING

By J. NISBET, MURIEL ADAMS AND JEAN ARTHUR

(Department of Education, University of Aberdeen)

SUMMARY. Short versions of Vernon's Arithmetic-Mathematics Test for pupils in the first and second years of secondary school were identified from the difficulty indices of items calculated from a sample of 100 pupils. In an independent sample, scores on ten 'best items' showed correlations over +0.90 with scores on the full 65-item test.

INTRODUCTION

Wood (1969) provides a summary of the rationale of tailored testing. In conventional testing, some items are superfluous in the scores of individuals because they are either too easy for the able pupils or too difficult for the weaker ones. In a tailored test, pupils begin with the same first item, but if their answer is right they are directed to a more difficult item, and if wrong, to an easier one. This happens with each response, so that the test 'zeros in' on a precise assessment of each individual. Mechanical devices, such as sticky tape and chemical pencils, were used in the development of the idea, but the technique can readily be handled by a computer with visual display.

Such procedures are unavailable as yet for large-scale screening or monitoring of standards, in which tests are applied to whole age groups in a region. A crude version of the idea, however, may be useful in reducing the time spent (by markers as well as by pupils) when the test used covers a wide range of ability, provided that the levels of difficulty of the items are known or have been established in a small-scale pilot run.

METHOD AND RESULTS

In an evaluation exercise in Grampian Region, Vernon's Graded Arithmetic-Mathematics Test (Senior) was applied to a one-in-four random sample of all pupils in the first and second years of secondary schools in the region. It was clear from marking the scripts that many of the items were largely a waste of time for certain of the pupils at this age level: the early items were too easy to discriminate effectively, and hardly any of the pupils were able to answer any of the items in the second half of the test. Indices of difficulty and discrimination* were calculated for a random sample of 100 scripts. The inefficiency (at this age level) of many of the items is demonstrated in Table 1, which summarises the indices for items at various stages through the test. The first ten and the last 31 items have poor discrimination indices (and therefore contribute little if anything to the test) because the first group is much too easy and the last group much too difficult.

Could we then mark only ten, or 20, of the items and obtain as valid a score as from the whole test of 65 items? When indices were calculated separately for the 51 pupils in first year of secondary school and for the 49 in second year, the ten best items for first-year pupils (items 13, 14, 17, 18, 19, 20, 21, 22, 24, 28) were different from the ten best for second-year pupils (items 12, 14, 16, 17, 18, 19, 24, 25, 29, 32). The term 'best items' has no independent validity, but is specific to the group and to the purpose for which the test is used. Our purpose differed from Vernon's and our pupils were from a more restricted age range than his standardisation group. It is not surprising therefore that our 'best items' are fewer and in a more restricted range than those selected by him in the process of test construction. It is fair to acknowledge that the test manual suggests omitting the early items when secondary school pupils are being tested. (This was not done in our survey because the testing was part of a screening for a remedial programme.)

To check the validity of our selection of items for a short version, it was necessary to draw a new random sample. It would clearly be invalid to check these 'best items' from the same set of scripts which had been used to identify the items. Consequently a fresh random

* The discrimination index was estimated by sorting the 100 scripts into a top third (33), a middle third (34) and a bottom third (33) on the basis of total score; then, for each item, counting the number in each third who answered correctly; then expressing the difference between the top third and the bottom third as a percentage (subtract, divide by 33, multiply by 100). This gives a rough estimate of the correlation between the item and the test total. The index of difficulty was simply the percentage answering the item correctly.

TABLE 1
DISTRIBUTION OF FACILITY INDICES AND DISCRIMINATION INDICES FOR ITEMS IN THE ARITHMETIC-MATHEMATICS TEST

	Items	Over 79	70-79	60-79	50-59	40-49	30-39	20-29	Below 20
Facility Index	1-10	9	1	—	—	—	—	—	—
	11-14	—	—	1	3	—	—	—	—
	15-24	—	—	4	2	1	2	1	—
	25-34	—	—	3	—	1	—	4	2
	35-65	—	—	—	—	—	—	1	30
Discrimination Index	1-10	—	—	—	—	1	2	2	5
	11-14	—	1	1	2	—	—	—	—
	15-24	2	3	2	2	1	—	—	—
	25-34	—	—	2	1	3	3	—	1
	35-65	—	—	—	—	—	3	5	23

sample of 100 scripts was drawn. There happened to be exactly 50 first-year pupils and 50 second-year pupils. Within the two samples of 50, the following scores were calculated:

- A. score on the total test;
- B. score on the 'best ten' items;
- C. score on 20 successive items, from no. 15 to no. 34.

The correlations of B with A, and C with A, given in Table 2, provide a check on the validity of the short versions. The correlations derived from the first sample are also given, for comparison. All the correlations are above +0.9, showing that the tailored versions are producing an order of merit as efficiently as the reliability coefficient of the whole test.

TABLE 2
CORRELATIONS BETWEEN SHORT VERSIONS AND WHOLE TEST

Correlation between total test score and—		2nd Sample	Initial Sample
First-year pupils: (N = 50)	best ten items	+0.91	+0.92
	items 15-34	+0.96	+0.96
Second-year pupils: (N = 50)	best ten items	+0.92	+0.90
	items 15-34	+0.96	+0.96

Discrimination indices were calculated for the 23 items involved in the scoring of the second sample as a check on the results obtained from the first sample. These showed some fluctuation, ten being within 0.09 of the previous figure, 12 varying by 0.10–0.19, and one differing by 0.21 from the previous figure.

We could certainly have saved the markers' time by instructing them to mark only ten of the 65 items. It does not follow that the same results would have been got if the pupils had been given only 10 or 20 items to answer, for the first ten easy questions may have helped some to settle to the task. For pupils of this age, the items after no. 34 could readily be ignored, though the effect of a time limit would be altered by changing the length of the test. To obtain equivalent norms for the short version for comparison with published norms for the whole test, is a relatively simple procedure, but would probably require a larger sample. There are other complications: since the item reduction is purely a statistical process, the content validity of the short test may be quite different from the original. Reliability will also be affected, though relatively slightly since most of the discarded items are irrelevant to position in the total score. If as few as ten items are used, many pupils will be grouped on the same score.

In time, the availability of item banks and computer facilities in schools may make these crude methods obsolete. For the present, however, they may be of interest to those who have

to use existing tests and hand scoring. Where large numbers of pupils are involved, selective scoring based on knowledge of item difficulty may sometimes be justified without serious loss of validity.

ACKNOWLEDGMENTS.—The assistance of Jennifer Welsh and Neil Ferguson, University of Aberdeen, and of Margaret Taylor and William Fordyce, Grampian Region, is gratefully acknowledged.

REFERENCE

WOOD, R. (1969). The efficacy of tailored testing. *Educ. Res.*, 11, 219-222.

(Manuscript received 4th February, 1980)

THE EFFECT OF HEADINGS ON CHILDREN'S RECALL FROM PROSE TEXT

By J. HARTLEY, JILLIAN KENELY, GWENDA OWEN AND M. TRUEMAN
(*Department of Psychology, University of Keele*)

SUMMARY. One hundred and seventy-five children studied a short passage in one of four conditions—control, passage with title, passage with headings in the form of statements, and passage with headings in the form of questions. The results showed that the passages with headings were recalled significantly better than the control passage or the passage with a title, on both immediate and long-term recall (14 days later). In addition, low ability children profited best from the text with headings in the form of questions.

INTRODUCTION

In guidelines for writing instructional text, it is common to find exhortations to use headings (e.g., Cooper, 1964; Wright and Barnard, 1975; Hartley, 1978; Hartley and Burnhill, 1978). It is unfortunate, however, that such exhortations are rarely supported by any evidence. Few studies have actually examined the effects of the presence or absence of headings, and whether or not headings in the form of questions are more effective than headings in the form of statements, and these few studies in themselves are not very encouraging (e.g., Robinson and Hall, 1941; Christensen and Stordahl, 1955; Klare *et al.*, 1958).

Although there is little direct evidence of the effectiveness of headings, there is indirect evidence which might lead us to suppose that the addition of headings (particularly in the form of questions) might be a useful instructional procedure. Work on titles, for instance, indicates that they can have an important role in helping the understanding and subsequent recall of obscure passages (Dooling and Lachman, 1971; Bransford and Johnson, 1972; Kozminsky, 1977). Similarly, work on questions suggests that questions asked before paragraphs of information lead to an increase in the recall of specific facts related to those questions, and that questions asked after paragraphs of information lead to an increase in the recall of more general material (e.g., Andre, 1979; Rickards, 1979).

In view of this research it might be expected that headings in the form of questions would lead to better recall than headings in the form of statements, and that the presence of both devices would lead to better recall than would their absence. To test out some of these notions we carried out the following enquiry.

METHOD

Participants

Approximately 200 second-year pupils from eight classes in a local comprehensive school took part. The pupils were streamed into bands of high, medium and low ability on the basis of primary school records and first-year examination results.

Materials

An extract (containing approximately 400 words) was adapted from a text by White (1963). This passage contained a description of the main events in the nursing career of Florence Nightingale, and was divided into six paragraphs of roughly equal length. Four typescript versions of the material were prepared:

- Condition A Control passage: no organisational aids.
- Condition B Passage with title.
- Condition C Passage without title, but with underlined headings before each paragraph. These headings were in the form of statements.
- Condition D Passage without title, but with underlined headings before each paragraph. These headings were in the form of questions.

The statement and the question headings corresponded to each other as far as possible (e.g., the statement 'The early life of Florence Nightingale' became 'What was Florence Nightingale's early life like?'). Each heading provided a close summary of the main

idea to be related in the following paragraph. The Flesch Reading Ease Score for each passage was 72, which suggests that it was suitable for these readers.

Method of Testing

A set of eight short-answer factual recall questions was constructed. The test measured the main points in the passage relating to Florence Nightingale's life and opportunities in nursing, etc. Because some questions contained more than one part, the total number of marks obtainable was 15.

Procedure

The children were allocated systematically—on an alphabetical basis—into the four treatment conditions in each of the eight classes. The experiment was conducted by the teachers in charge of each class following instructions provided by the authors. These were:

"The purpose of this investigation is to find out whether there are any effects on the children's recall of the passages as a result of how the passages are organised. Would you please hand out the passages (each pupil's name is on the top of the sheet) and allow your pupils to read them. When each pupil has finished reading the passage, would you please collect them in and then hand out the answer sheets which are also headed with the pupil's names. *Please do not give out the answer sheets until all the passages have been collected in.* There is no set time limit, but the experiment shouldn't take more than 40 minutes to complete. Thank you for your co-operation."

In fact, all the groups finished within 30 minutes.

Two weeks later, without warning, the pupils were re-tested. They did not read the passage again before doing the retention test. The data, analysed below, are from those 175 pupils who were present for both test sessions.

RESULTS

TABLE 1

THE AVERAGE SCORES (OUT OF 15) OBTAINED BY PUPILS IN THE THREE ABILITY GROUPS IN THE FOUR CONDITIONS

	Condition A (Control)	Condition B (Title)	Condition C (Headings as Statements)	Condition D (Headings as Questions)
High ability	7.5 (N = 17)	7.9 (N = 16)	9.3 (N = 18)	9.1 (N = 18)
Medium ability	6.1 (N = 19)	5.0 (N = 20)	6.9 (N = 18)	6.3 (N = 17)
Low ability	1.9 (N = 10)	2.9 (N = 6)	3.5 (N = 7)	6.9 (N = 9)
Overall	5.7 (N = 46)	5.8 (N = 42)	7.4 (N = 43)	7.6 (N = 44)

The overall results are shown in Table 1. To test the significance of these results we carried out four-way analysis of variance, with one repeated measure (conditions $\times$ ability $\times$ sex $\times$ test-retest). Three main effects and two interactions emerged as statistically significant. These were:

(i) There was a significant difference between the average test (7.6) and retest scores (5.7) ($F = 117.49$, $df\ 1,151$, $P < 0.001$). Thus there was some forgetting over the two-week period. However, because there were no significant interactions between the test scores and the other variables it can be assumed that this forgetting was a consistent factor throughout the data. Therefore the test and retest scores are averaged and not considered separately in the remaining analyses.

(ii) There was a significant difference between the four treatment conditions ($F = 7.56$,

G

df 3,151, $P < 0.001$). The overall averages were 5.7 for Condition A, 5.8 for Condition B, 7.4 for Condition C and 7.6 for Condition D. Further analyses by Scheffé tests showed that there were no significant differences between the scores in Conditions A and B, and between those in conditions C and D. However, the combined scores of Conditions C and D were significantly superior to the combined scores of Conditions A and B ($F = 4.88$, df 3,171, $P < 0.01$).

(iii) There was a significant ability effect ($F = 45.43$, df 2,151, $P < 0.001$). The high ability pupils scored on average 8.5, the medium ability 6.1 and the low ability 3.8 respectively. Scheffé tests showed that these differences between the groups were all significant at $P < 0.01$.

(iv) There was a significant interaction between conditions and ability ($F = 2.47$, df 6,151, $P < 0.05$). This interaction is apparent in Table 1. It can be seen that Condition D had a greater effect for the low ability group: these pupils did as well as medium ability pupils in this condition.

(v) Although there was no overall sex difference, there was a significant sex $\times$ ability interaction ($F = 3.69$, df 2,151, $P < 0.05$). High ability girls did better than high ability boys, but there were no significant differences in this respect for the medium and low ability pupils.

DISCUSSION

These results are clear-cut. The effect of giving the passage a title was negligible. The effect of interspersing headings in the text was striking. Overall it made no difference whether or not these headings were in the form of statements or questions, but low ability pupils did significantly better with headings in the form of questions.

These results lead us to make the following observations:

(1) The headings had a marked effect—and this occurred in prose which is not normally considered suitable for headings. (Headings generally appear in more technical writing than in English literature.) The findings suggest the value of including headings in light non-fiction as well as in technical prose.

(2) The fact that lower ability children did better with sub-headings in the form of questions is intriguing but requires further study. Our findings are in line with those of Rickards and Hatcher (1978) who found that the performance of poor comprehenders could be greatly enhanced by placing questions after paragraphs in text but that the effects were not so marked for more able readers. These findings, together with our own, suggest that even if one does not always do it, it will always be worth considering the use of headings in the form of questions when writing instructional materials.

Our results suggest that it would be of value to replicate this kind of study with different passages and children of different ages. Investigators are beginning to realise that learning to read instructional text involves learning processes in addition to those of learning to read (e.g., Brown and Campione, 1978). At present it is not clear when children begin to acquire the methods and techniques of understanding that adults use with this kind of text. It has been suggested, for example, that underlining in text can hinder the comprehension of young children (cf Rickards and Denner, 1979 with Hartley *et al.*, 1980) and that the use of multiple typographic cues is confusing for them (Hershberger and Terry, 1965).

Certainly, as far as future research is concerned, if our findings can be replicated in different situations, then it should be possible to delineate whether or not some kinds of questions and some kinds of text are more critical than others.

REFERENCES

- ANDRE, T. (1979). Does answering higher-level questions while reading facilitate productive learning? *Rev. educ. Res.*, 49, 280-318.
- BRANSFORD, J. D., and JOHNSON, M. K. (1972). Contextual pre-requisites for understanding: some investigations of comprehension and recall. *J. verb Learn. verb. Behav.*, 11, 717-726.
- BROWN, A. L., and CAMPIONE, J. C. (1978). The effects of knowledge and experience on the formation of retrieval plans for studying from texts. In GRUNBERG, M. M., MORRIS, P. E., and SYKES, R. N. (Eds.), *Practical Aspects of Memory*. London: Academic Press.
- CHRISTENSEN, C. M., and STORDAHL, K. E. (1955). The effects of organisational aids on comprehension and retention. *J. educ. Psychol.*, 46, 65-74.
- COOPER, B. N. (1964). *Writing Technical Reports*. Harmondsworth: Penguin.

- DOOLING, D. J., and LACHMAN, R. (1971). Effects of comprehension on the retention of prose. *J. exp. Psychol.*, 88, 216-222.
- HARTLEY, J. (1978). *Designing Instructional Text*. London: Kogan Page; New York: Nichols.
- HARTLEY, J., BARTLETT, S., and BRANTHWAITE, J. A. (1980). Underlining can make a difference—sometimes. *J. educ. Res.*, 73, 218-224.
- HARTLEY, J., and BURHNILL, P. (1978). Fifty guidelines for improving instructional text. In HARTLEY, J., and DAVIES, I. K. (Eds.), *Contributions to an Educational Technology, Vol. 2*. London: Kogan Page; New York: Nichols.
- HERSHBERGER, W. A., and TERRY, D. F. (1965). Typographic cueing in conventional and programmed texts. *J. appl. Psychol.*, 49, 55-60.
- KLARE, G. R., SHUFORD, E. H., and NICHOLS, W. H. (1958). The relation of format organisation to learning. *Educ. Res. Bull.*, 37, 39-45.
- KOZMINSKY, E. (1977). Altering comprehension: the effect of biasing titles on text comprehension. *Memory & Cognition*, 5, 4, 482-490.
- RICKARDS, J. P. (1979). Adjunct post-questions in text: a critical review of methods and processes. *Rev. educ. Res.*, 49, 181-196.
- RICKARDS, J. P., and DENNER, P. R. (1979). Depressive effects of underlining and adjunct questions on children's recall of text. *Instruct. Sci.*, 8, 81-90.
- RICKARDS, J. P., and HATCHER, C. W. (1978). Interspersed meaningful learning questions as semantic cues for poor comprehenders. *Reading Res. Q.*, 13, 538-553.
- ROBINSON, F. P., and HALL, P. (1941). Studies of higher-level reading abilities. *J. educ. Psychol.*, 32, 241-252.
- WHITE, E. E. (1963). *Everyday English, Book 2*. London: Dent.
- WRIGHT, P., and BARNARD, P. (1975). Just fill in this form: a review for designers. In HARTLEY, J. (Ed.) (1980), *The Psychology of Written Communication: Selected Readings*. London: Kogan Page; New York: Nichols.

(Manuscript received 31st October, 1979)

BOOK REVIEWS

GALTON, M., SIMON, B., and CROLL, P. (1980), *Inside the Primary Classroom*. Oxon: Routledge and Kegan Paul, pp. v+203, c. £8.95, p. £4.95

The book claims to report the first large scale observation study of primary classrooms. It comprises observations on an opportunity sample of 19 schools in three LEAs, although some, for example the 10-year-olds, are drawn from only one LEA. The 58 teachers and their pupils were observed on occasion through the year during which time the teachers were each observed for a little under six hours in total, and selected pupils for a little under 11 hours.

Concern about the validity of the picture gained by such a thin observational spread is exacerbated when the content of the observation is considered. The focus was on 'normal teaching and learning activities' and it was thus surprising to note that activities such as PE, music, drama, singing and watching TV were not included in this definition and thus not observed. This is of importance since it serves to distort the curriculum balance portrayed. For instance the figure of 65 per cent of time spent on maths and language is a proportion much higher than other studies here reported. Their finding that 'primary schools concentrate very heavily on basic number and language work' thus needs to be tempered with the realisation that the data are not based on the whole curriculum.

There follows description of the 'typical' pupil based on the averages of all pupils irrespective of age, prior to cluster analyses of pupils and teachers to create types. Teacher types are based on interaction data which disregard classroom and curriculum organisation and four (split into six) types emerge. Four pupil types are also apparent but the decision not to include such variables as age, sex, or ability makes interpretation, and comparison with other typologies, difficult. There is an argument put forward that the proportion of the different pupil types in a class could relate to teaching approach but the analysis provided is not based on change data and can only be regarded as very tentative at this stage.

In many ways this is a rather curious text following no coherent pattern. There is no systematic review of other studies in this area, the first chapter comprising a self-justificatory attack on questionnaires, and thereafter chapters reporting findings are interspersed with chapters on the history of primary education and descriptions of the Plowden ideology. Indeed this ideology provides a thread which runs throughout the book whereby there is a continuing urge to compare prescription with practice despite the fact that the evidence provided by their data is incongruous with evidence needed to test the prescription.

In spite of the limitations outlined earlier the authors are on stronger ground when limiting themselves to descriptions of practice, and here most of their findings are supportive of earlier studies. In these three LEAs 'pure' progressive teaching is rare, teachers tend not to ask many open or high level questions and classroom groups are set up for organisational rather than pedagogic purposes in which the sexes tend not to interact.

The next volume will report the process-product findings and will be of greater interest. Nevertheless the validity of those results will depend on the quality of the process data reported in this volume. The limitations apparent will thus need to be borne in mind.

NEVILLE BENNETT.

JENSEN, A. R. (1980). *Bias in Mental Testing*. London: Methuen, pp. vii+786, c. £15.00.

"An outstanding contribution to mental testing" as R. E. Rawles described it, or in the view of Leon Kamin, "a dud that simply fails to ignite". The difference in judgment between these two reviewers merely reflects the confusion in aims of this strange book.

On the face of it the book is a technical treatise: a detailed consideration of the possibilities of bias in mental tests leading to a set of prescriptions for their use in education and guidance. Actually it is a laborious 800-page further defence of the tired thesis that mental tests all share a single common factor, the genetically based general intelligence, 'g', and that differences between ethnic groups and between social groups in mental test performance cannot be explained away in terms of the cultural bias of the tests. The 15 points IQ difference between US blacks and whites must be attributed in large part to 'g'; ergo blacks on average are less well-endowed genetically with intelligence than whites.

Although the latter conclusion is undoubtedly the main one to which the book leads, Jensen, however, prefers to sidestep it, dropping in only the occasional reminder that it is coming: "It would be difficult to make the case for the hypothesis that heritability of IQ is less than 0.50 (p. 244)"; "most published estimates of IQ heritability fall in the range 0.6 to 0.9" (p. 427) and the corollary, "the race differences observed in drawing and copying tests are more parsimoniously interpreted as differences in rates of mental maturation than as cultural differences" (p. 713). His last chapter is devoted instead to a discussion of the abuses of mental tests in education; and his own proposals for the proper use of them are remarkably modest and humane in the light of the fiery segregationist policies attributed to him by his publishers. Who would quibble with: "The great amount of overlap of . . . (the black/white IQ) . . . frequency distributions absolutely contradicts the racist philosophy that persons of different races should be treated differently, one and all, only by reason of their racial origin." (p. 738).

Whether there is a scientifically legitimate means of estimating IQ heritability and if there is what the estimate should be has been argued about ad nauseam elsewhere. Jensen softens his earlier position on it slightly, conceding that it is "rarely the case that the data meet all the assumptions underlying any one method" (p. 244), but then decides to avoid this issue altogether "because even an elementary explication of heritability analysis is beyond the scope of this book" (p. 427). It needs to be said that any estimation method applied to human beings is incapable of meeting the basic assumptions of Fisher's heritability estimation procedure, the assignment of genetically similar individuals at birth to random environments (Kempthorne, 1978; Vetta, 1980). Moreover, in one of the most recent attempts, within the tradition that believes it can be done, the large-scale Texas adoption study (Horn *et al.*, 1979), the conclusion is reached that: "A simple polygenic model featuring an extremely high heritability of IQ would seem very difficult to reconcile with these data . . .".

Heritability estimation is of course the most direct means of determining whether 'g' an "individually adaptive behavioural capacity" is located in the genes as the eugenics theory Jensen espouses predicts that it should be. The other method with which the first two-thirds (10 chapters) of the book is concerned is validation in the psychometric sense. Via the early history of intelligence theorising and measurement in Galton's laboratory, through the multitudinous data collected in the construction and application of mental tests and the investigation of their correlates, Jensen sets out the evidence that Spearman's original theory of one single factor of general intelligence is essentially correct. This trait which is most strongly present in non-verbal tests concerned with, in Spearman's original terms, the "education of relations and correlates", is normally distributed in the population and hence is the product of multiple determination from the genes; has certain physical correlates, e.g. head size (!); and predicts subsequent success in school and adult life. To me, though, it is the hiccups in this plausible story that are of particular fascination. For example, it is ironic that Francis Galton and James McKeen Cattell in their first attempt to operationalise the theory produced strings of psycho-physical measurements which failed to inter-correlate with each other and with any measure of educational success. It was only with the advent of Binet, who designed his test specifically for educational selection purposes, that measures which are essentially concerned with the attainment of cognitive skill as opposed to biologically fixed characteristics were available to operationalise the intelligence construct. Ironically, although these tests, as we might expect, tend to predict subsequent attainment reasonably well (if 0.50, with US high school achievement is reasonably well) their success in predicting later measures of, for example, occupational competence is less impressive; correlations quoted here (p. 348) are in the order of 0.22, i.e. 4 per cent of the variance explained.

There is also conflicting evidence about the direction of the correlation between IQ and some physical measurements. Thus although brain size, stature and basic metabolic rate show modest positive correlations with IQ, for others like reaction time and 'EEG' the correlation may be zero or reversed.

The main evidence for the existence of 'g' comes from factor analysis, the technique which Jensen uses to provide evidence in support of his theory, without acknowledging the point that in the form in which he uses it it is incapable of doing anything else. Setting aside the controversy in the 1940s between the followers of Thurstone, who supported the 'group

factor' model of intelligence, and the followers of Spearman who supported the 'general factor' one, *exploratory* factor analysis can produce no structure of mental ability other than some kind of hierarchy. The possibility of such structures as 'the simplex', 'circumplex', or 'radex' as described by Guttman, for example (Shye, 1978) in which one ability feeds another can never be tested by it. For this purpose confirmatory methods, such as those developed by Jöreskog and Sörbom (1979), are required in which a structural model is specified and its fit tested against the data. It is ironic that Jensen in presenting a matrix of correlations between mental test performances over age (p. 279), describes the structure as a simplex, a structure that his trusted (exploratory) factor analysis would never allow him to discover. It seems worth repeating the old adage that you get out of factor analysis what you put into it; in this case Spearman's theory of intelligence.

It is not until chapter 10 that Jensen gets on to the issue of test bias, which he defines in two ways. The first definition identifies bias basically with changes in the predictive validity of the test from one sub-group of a population to the next; the second is connected with the test's construct validity, i.e. bias exists when the factor structure of the test (item characteristics, factor loadings on 'g', etc.) differ between sub-groups. What cannot be claimed as bias is the mere existence of differences in mean score between sub-groups on a particular test. These differences have to be explained and one of their causes may be test bias; but in themselves they say nothing about bias. What follows are four chapters containing by far the most interesting material in the book, the evidence that Jensen has collected (four filing cabinets full) to refute the view of his earlier critics that IQ tests are biased against blacks and working-class groups.

The evidence here is impressive and one has to concede Jensen's case that whatever it is the tests measure is on average in greater abundance among whites than among blacks. And if the skills involved are worth having then clearly educationists do have an important remedial job to do. The problem is, though, that Jensen is still pursuing his basic thesis that the test differences between ethnic groups are innate, so he tends to overlook some of the evidence which could help teachers to see where the origins of the black children's difficulties might lie. Thus when we learn on page 592 that nine IQ points can be attributed to coaching and practice between *two* testing occasions, the 15 IQ point difference between blacks and whites tends to look less impressive. Moreover figure 12.1 on page 595 shows that blacks gain significantly more in IQ from practice than do whites. Yet nowhere does Jensen consider the possibility that in aggregate this small interaction effect and the myriad others arising from opportunities to practise reasoning skills, which may or may not occur in a child's early life, could account for the black-white differences. Jensen's technique is to work through a formidable list of influences, taken two or three at a time, which have been postulated to cause the IQ difference, e.g. coaching, practice, race, language and sex of tester, test anxiety, self-esteem etc. All described as possible 'external sources of bias'. Because most of the interaction effects that he finds between race and these influences on IQ are small and sometimes statistically insignificant, he considers they can be discounted. Nowhere does he consider the possibility that in aggregate their effect on the IQ difference could be substantial. A single multiple regression analysis with interaction terms included of data collected in a longitudinal study of all these postulated influences is the obvious way to examine this possibility, but nowhere does Jensen appear to consider it.

By any reckoning, despite the inordinate length of his book, Jensen fails to present a convincing case for the theory that he is espousing. This still leaves us with the issue of what if any place mental tests do have in education. Jensen, though opposed to their use routinely for selection purposes in primary education, does support their use for college entrance. He maintains that school leaving certificates in the USA are little more than records of attendance, whereas the college aptitude tests are a proper test of the potential students' abilities. In this country the public examinations, GCE and CSE, do reflect attainments in the secondary school syllabus and despite their imperfections have always been taken to be the fairest way of deciding who should go on to higher education and who should not. But this viewpoint may be too complacent. On the other hand, earlier on in the school career there is obviously a need for evidence of weaknesses in children's development of cognitive skills. Thus in educational attempts to remedy such deficiencies (e.g. in compensatory education, for example) it would seem that mental tests to monitor improvements do have an important part to play. The pity of Jensen's book is that the examination of evidence relating to this kind of issue is compressed into a mere five chapters.

In rushing to the defence of Jensen's 'right to know', many psychologists have overlooked the point that racial differences in IQ were not the real issue. The main thesis of his famous *Harvard Educational Review* article (Jensen, 1969) was that attempts to modify what, in his terms, were children's relatively fixed cognitive capacities were doomed to failure. He even went so far as to advance a theory (ignored in this book) of 'two types' of intelligence to explain it. As Hunt (1969) argues, at this relatively primitive stage of our understanding of the complex interactions between genes and environment through which human intelligence is formed, we can ill afford such pessimism about the possibilities for its growth. The research stance that uses tests to isolate what is 'fixed' as opposed to what is 'modifiable' underwrites an educational technology which will inevitably be seen as repressive rather than liberating. This denies the vitally important role that psychometrics can have and should have in providing the most rational means we have available for monitoring human growth. If cast in a different mould Jensen's book could have made an important contribution to it. Let us hope that now he has got the desire to refute his critics 'out of his system' he will concentrate his efforts in this educational direction instead.

JOHN BYNNER.

REFERENCES

- HORN, J. M., LOEHLIN, J. C., and WILLERMAN, L. (1978). Intellectual resemblance among adoptive and biological relatives: the Texas Adoption Project. *Behav. Genet.*, 9, 177-208.
 HUNT, J. MC.V. (1969). Has compensatory education failed? *Harvard educ. Rev.*, 39, 449-483.
 JENSEN, A. R. (1969). How can we boost IQ and scholastic achievement? *Harvard educ. Rev.*, 39, 1-123.
 JORESKOG, K. G., and SÖRBOM, D. (1979). *Advances in Factor Analysis and Structural Equation Models*. (Edit. J. Magidson). Cambridge, Mass: Abt Associates.
 KEMPTHORNE, O. (1978). Logical, epistemological and statistical aspects of nature-nurture data interpretation. *Biometrics*, 34, 1-23.
 SHYE, S. (1978). *Theory Construction and Data Analysis in the Behavioural Sciences: A Volume in Honour of Louis Guttman*. San Francisco: Jossey Bass.
 VETTA, A. (1980). Concepts and issues in the IQ debate. *Bull. Brit. Psychol. Soc.* (In press).

KITWOOD, T. (1980). *Disclosures to a Stranger: Adolescent Values in an Advanced Industrial Society*. London: Routledge and Kegan Paul, pp. 291, c. £11.50.

Every research method has its attendant problems. Dr Kitwood eschews the 'hard' empirical line in favour of in-depth interviews with samples of adolescents, and the result is a collection of themes examined and illustrated via the products of these interviews.

The problems with such methods are well attested: establishing the reliability of the statements, determining the extent to which the situation and the interviewer constrain or transmute the respondents' statements, providing a report which is as fair a reflection of the corpus of data as is reasonable. The author is well aware of the problems, indeed he devotes 40-odd pages to them and then raises the issues again in subsequent chapters. Whilst this is laudable in one sense, it is somewhat irritating in another. This reviewer confesses to having felt that he would never reach the research itself.

Data were obtained by inviting the teenagers to discuss a selection of issues which was made from an array of 15 statements, a procedure which Kitwood has described in earlier publications. To offer such choice has clear benefits in avoiding interviewees straining to invent responses to irrelevant probes. Its problem, of course, is that it may well eliminate other types of potential response such as the taken-for-granted issue, or the too-private-to-talk-about topic. Thus one might choose to believe that teenagers are not particularly concerned with sexual matters, as this study suggests; or one might feel that there are other reasons for its exclusion.

A general survey of the adolescents' responses in terms of topic choice revealed the popularity of issues which permitted a favourable projection of the self "when you got on really well with people", and a surprising unpopularity of issues relating to being misunderstood or being socially isolated. As the author points out there is here an interesting reflection of one dominant motive behind the choices.

Analyses of topic selected by sex, age and social class are also presented in some detail, but as numbers are necessarily very small there is little to be made of these other than the general impression that there appear to be some differences in topic selection by social class.

The latter part of the analysis, and the major portion of the research, consists of detailed studies of four topics: family life, peer relationships, formal and informal work, and the development of 'self-values'. Again, numbers of respondents in analysis by age, social class, and sex are small, but there is a rich exploration of the social contexts of these urban teenagers. And if a classification of seven types of 'conflict in the home' suggests that there could be only one or two manifestations of each, this is probably as it should be. The study is fine-grained, and a useful antidote to psychological studies of the past which might have implied that adolescents were cloned.

Tom Kitwood's appraisal which concludes the book is honest and clear, if not world-shattering. He scrutinises his methods (yet again), and summarises his principal finding as the adolescents' "strong concern about competence in the interpersonal domain" (p. 279). This seems to corroborate the impressions one gains from eavesdropping in coffee bars and reading the problem page in *Jackie*.

For anyone working with adolescents or teaching courses on adolescence there is much to be gained from the book, though it is not easy to read. Apart from the 'soul-searching' on methods, there are large tracts of transcript, replete with all the original 'ums' and 'ers'. These are no doubt to lend authenticity, but they soon begin to pall.

GEOFF BROWN.

KLAUSMEIER, H. J. and ASSOCIATES (1979). *Cognitive Learning and Development: Information-Processing and Piagetian Perspectives*. Cambridge, Mass.: Ballinger, pp. v+370, n.p.

The bulk of this book is by Herbert J. Klausmeier, describing his own studies of his conceptual learning and development theory, which he refers to as CLD theory, and its relation to what he takes to be Piagetian theory. Klausmeier's contributions are bolstered by chapters by Brainerd, Hooper *et al.*, and Murray, which describe closely related studies and provide more wide-ranging discussion. This hybrid format is an interesting idea, though I do not regard it as successful in this case. It has been assembled in a way that seems hurried, even to the extent that different titles are given on the book's title page and the dust jacket. There is much repetition, as if each of the chapters had been written as a separate report and they were then strung together. This impression is strengthened by the lack of logical sequencing in the presentation of information.

The book tries to do two things. The first is to show that CLD theory is fundamentally different from Piagetian theory. The second is to investigate the interrelationship of various indices of development in terms of CLD theory, Piagetian theory and educational attainment.

I am not personally much impressed by CLD theory. It is claimed to be an information-processing theory. This term, as Klausmeier rightly points out in the preface, has now been used in so many different ways that it is hard to see what it means. In saying that he takes it to mean that cognitive development can be explained in terms of acquiring, organising, storing and retrieving information Klausmeier is hardly giving much away.

The nub of CLD theory turns out to be that concepts are developed through four levels, which I will illuminate with reference to the example of face recognition used by Klausmeier. At the concrete level, recognition of an instance is only possible when it occurs in the same form in the same context, so a face would only be recognised in the same pose and physical location. At the level of identity, instances are recognised despite changes in form or context, so that the recognition of particular people would be possible. At the classificatory level instances can be generalised across particular forms and contexts, so that a concept of faces or persons as a class of stimuli is formed. At the formal level, the concept is understood sufficiently well for a socially acceptable definition to be provided.

There are compelling theoretical and empirical objections to a theory of this type. Firstly, it is not at all clear why the concrete or identity levels should be regarded as involving concepts. In many senses they are pre-conceptual. Secondly, the levels differ in the complexity of the skills demanded, so that it is hardly surprising when Klausmeier 'discovers' that children can recognise the shape of equilateral triangles (identity level of concept of equilateral triangle) before they can explain that all three sides of an equilateral triangle are of the same length (formal level). The use of face recognition as an illustrative example is particularly unfortunate, since there is no reason to believe that even young infants are ever entirely at the concrete level. Moreover, it is claimed (p. 13) that adults remember faces better than children *because* they have a concept of a person at the classificatory and formal levels. Despite this claim, most studies have shown that children are quite good at remembering faces, and it is easy to ascertain that they know what a person is.

The idea that knowledge is represented at different levels during human development is, of course, a cornerstone of Piagetian theory, though Piaget's view as to what the levels actually are is rather different. Klausmeier, however, persistently asserts that CLD theory is fundamentally different from Piagetian theory. His reasons why this is the case are astonishing. On page 5 we are told that in CLD theory development is seen as an *interaction* of learning and maturation, and that what the individual is capable of learning is related to the mental operations that have emerged and are functional. Yet this is exactly what Piaget has always maintained, as Brainerd actually explains (on p. 227) in this volume.

My own view, then, would be that CLD theory has none of the advantages of Piagetian theory and all (and more) of its disadvantages.

Turning to the empirical content of the volume, matters are no better. The principal method used is to assemble batteries of CLD measures, Piagetian measures, and educational attainment measures. Children's performance on these measures is then 'analysed' by intercorrelation and factor analysis. The results are as might be expected. Most of the measures correlate with most of the other measures. There is an abundance of data and no real findings. Although Brainerd and Hooper seem to realise that all is not quite as it should be, no one really questions whether a method like this could ever be capable of deciding between one theory and another. Nor is the validity of attaching numerical labels to the skills measured on Piagetian tasks discussed. There are no attempts to test between the theories using the more convincing method of deriving different predictions and conducting appropriate experiments.

Finally, it is irritating to find that many of the claims made in the book cannot readily be checked because over forty of the references in the bibliography are to unpublished work.

ANDREW YOUNG.

LEEMING, K., SWANN, W., COUPE, J., and MITTLER, P. (1979). *Teaching Language and Communication to the Mentally Handicapped*. Schools Council Curriculum Bulletin 8. London: Evans/Methuen Educational. pp. xiii + 400, c. £9.25.

Teachers and psychologists alike, when faced with the practical problems of planning intervention aimed at improving an ESN(S) child's language, are only too aware of the inadequacies both of existing language assessment techniques and sources of information about teaching approaches.

This book describes a Schools Council project which has attempted to tackle these problems. Its brief was to survey existing language teaching practices and attitudes to language teaching in ESN(S) schools, to identify good practice, and to identify and evaluate teaching materials.

The survey of the communication skills (production, reception, articulation and, interestingly, use of gesture) of 1400 ESN(S) children in the North-West, resulted in some disturbing, if not entirely unexpected, findings. For example, 42.5 per cent of those aged 16 years or over could not use grammatical constructions; a quarter of the sample had no

identifiable comprehension; a third were at or below the level of imitating single words. There is no reason to suppose that the sample is not representative of the national population.

A summary of the stated aims of teachers in ESN(S) schools and an examination of what happened in practice revealed very little relationship between the two. The emphasis placed by the authors on the importance of language in the curriculum and the proposal of a detailed teaching model, which is predominantly behavioural, may help to reduce this gap. The model is not entirely new but the detailed accounts of practical teaching sessions which illustrate the way in which theory can be translated into practice is excellent. After a chapter in which, perhaps a little belatedly, the communication process is defined, four successive chapters give more examples of the use of systematic observation; vocabulary teaching; non-verbal communication; the use of questions and a review of four published language curriculum packages.

Although the order in which material is presented in the earlier part of the book makes following the line of argument a little difficult at times, it is fair to say that the authors have not only met the requirements of their brief but have also provided a wealth of tried, exemplified teaching ideas cogently related to theory. This is a book which should provide both teacher and psychologist with a basis for discussions about curriculum planning and which any teacher of ESN(S) children will need to study and use if the much-needed improvement in the quality of language teaching is to occur.

IAN BERRY.

O'NEIL, H. F. (Jr.) and SPIELBERGER, C. D. (Eds.) (1979). *Cognitive and Affective Learning Strategies*. New York: Academic Press pp. v+327 c. \$21.50.

There is no doubt that many people who undertake courses of study, or who study in private, do not derive as much benefit as they might because they approach what they are doing in an inefficient or inappropriate way. For this reason the topic of learning strategies is of much interest.

O'Neil and Spielberger's edited book sets out to deal with this topic in three sections: cognitive approaches to learning strategies, effective approaches to learning strategies, and alternative approaches to learning strategies. Within each section three chapters dealing with separate aspects are followed by a fourth chapter that tries to effect some kind of overview.

This is a reasonable plan, and there is much of interest in the book. Despite this, however, the book fails to be much more than a collection of individual chapters. Some of these chapters are good, some not so good.

There are several reasons why the individual chapters do not add up to a proper treatment of the topic. In the first place, the stated plan is not followed. Section I (cognitive approaches to learning strategies) is sound, but Section II (affective approaches to learning strategies) is poorly organised. The chapters by Judd *et al.* and Rigney *et al.* seem to be of more relevance to Section I. Spielberger *et al.* discuss test anxiety, an affective variable of obvious importance, but one that would not seem to have much direct relevance to learning strategies. Not surprisingly, in providing the overview to the chapters in this section, Sarason is rather obviously disconcerted by their disparate nature. The chapters in Section III (alternative approaches to learning strategies) are also heterogeneous.

A further difficulty arises from the fact that the emphasis in many of the chapters is not on explaining what learning strategies are available, but on attempting to evaluate how successful some of the different strategies are. In effect, this makes the book useless to anyone whose interest is in applying strategies, either as learner or teacher, since the various strategies are typically only described in a brief and abstract manner. Moreover, the success of strategies is usually evaluated in terms of a comparison of test scores of a group of people instructed to use a particular strategy with test scores of control groups not given any instruction as to strategy. This approach is much too narrow to be given such a central place as an evaluative 'strategy'. It assumes that people will derive full benefit from a strategy without having had much experience of using it. It ignores the fact that people who are not using the strategy investigated make use of alternative strategies of their own devising. It does not

adequately take into account the possibility that different strategies may suit different people, or different purposes. Although the authors of most of the chapters seem to be well aware of such difficulties, they are often passed over rather quickly. The important question as to why people who know the most effective strategies often cannot be bothered to use them is also set aside.

The book suffers from the apparent lack of external control often found in books that report data. Many of the studies described would be considered too preliminary or inconsequential for publication in a scientific journal. Some of the better studies have already been published separately.

I think, then, that this is a book that will mainly be of interest to people engaged in research similar to that of the authors. Although this is perhaps all that was intended by the editors, it need not have precluded a more thorough and satisfying examination of the topic of learning strategies, and a splendid opportunity would seem to have been missed.

ANDREW YOUNG.

RAGGETT, M., TUTT, C., and RAGGETT, P. (Eds.) (1979). *Assessment and Testing of Reading. Problems and Practices*. London: Ward Lock Educational, pp. 207, £4.95.

This book has been produced in response to an awareness among educationalists that "unless the profession was able to provide informed comment on reading tests and assessment techniques, the voices of practising teachers might not be heard in important matters of educational policy concerning national, regional, local education authority and school-based monitoring," (p. 25). Among the skills likely to command this authority is the power to test reading, for this is also "the power to determine the reading curriculum". To gain that skill and power demands a very adequate understanding of models of reading, the uses, types and characteristics of tests, the meaning of test scores and the interpretation of results and their relevance to decision-making. In addition to these requirements teachers should be aware of other methods of assessment, informal testing, informal inventories, and other ongoing practices which are being carried out by teachers, educational psychologists and others both in ordinary classrooms and special centres. To provide a good grounding in such knowledge and highlight issues in this area of reading is then a major aim.

Although literature in the field of reading has been extensive, the editors believe that special objectives will be achieved by the production of another book. Arising out of their own recent involvements in reading research they became "acutely aware of the lack of any up-to-date book readily available to teachers, which acted as a bridge between those pioneering the development of new thinking in the centres of research and the teachers in the field." This book is designed to be that bridge. In examining its content, structure and style of presentation we should bear in mind these dual aims. We can then expect it to contain much new material and the presentation to be clear and informative.

In the Introduction the editors mention the measures adopted to ensure an integrated coverage by the sixteen contributors. They also state that, in the interests of greater understanding they "have tried to develop basic issues in spiral form". It is a little unclear as to what they mean by this. It could mean that the same topics have been presented throughout but at successive levels of difficulty, or that the easier topics precede the more difficult ones. It can hardly mean the latter, since the book opens with the very difficult subject of "What is reading? Which model?" Although Elizabeth Goodacre develops this in her usual clear style and the models—those of the researcher, the test constructor, the teacher—are simple, yet it is impossible for her to avoid the presentation of more difficult material, not made easier by the necessity of compressing this in summary form. The teachers who read this section (a midget sample of putative readers) did find it daunting.

An apparent need to conserve space, more noticeable in the first sections of the book, that is, the theoretical as distinct from the sections presenting practical measures, has led to too great an emphasis on references to back up statements—in one instance thirteen are given at the end of a short paragraph. And possibly to statements which are either incomplete or

ambiguous. For instance, "A score on one test cannot be directly related to a score on another test, which may not be measuring the same thing. (To use an analogy, measuring a child's height will give little indication of his weight.)" (Introduction). One contributor, not writing specifically on the cloze test, observes that the cloze procedure is one "whereby a text is taken, with no alteration and words are deleted at random—say every sixth word." One feels too that a link has not been uncovered in the following statement: "Broader concepts of measurement such as reading ages . . . may tell one that pupil A is 'better' than pupil B. Why he is better is left a mystery. A possible answer to these problems, which Bruce Choppin discusses, is the use of the Rasch formula for measuring each item in a bank of test items."

Teachers will appreciate the more practical parts of the book, informal inventories, informal testing, individual and group assessment, descriptions of the Barking experiment and the Aston Index. It is a pity that the reference to multiracial test items in Helen Quigley's section on the London Test should be contained in a single sentence.

Finally, what is new? How new is new? Much of this book is not, either in thinking or practice. But it may be new to the teachers who read it. A lot of work has gone into it, but I fear it cannot be regarded as a first-rate contribution to the literature.

GERTRUDE KEIR.

RYRIE, A. C., FURST, A., and LAUDER, M. (1979). *Choices and Chances*. London: Hodder and Stoughton for the Scottish Council for Research in Education. pp. i+140, n.p.

This volume represents the first fruits of a longitudinal research project examining the ways in which young people progress through secondary schools in Scotland and into work or continued education. This report (others are anticipated) is specifically concerned with the subject choice of pupils at age 13 years or thereabouts and with their future educational and occupational intentions.

The sample of pupils interviewed consisted of 74 boys and 74 girls in each of eight comprehensive schools in two areas of Scotland, one rural and the other industrial. Within each school the pupil sample was split evenly between two successive year groups so, in effect, two cohorts of pupils were followed through until they left school. All pupils participating in the research (nearly 1,200 in all) were interviewed at various stages on their way through secondary school. Teachers and parents also participated in the research.

Choices and Chances is almost exclusively concerned with reporting the findings of these interviews. Very little attention is given either to relating this project's research results to those of earlier studies, or even to an evaluation of other research work on subject choice or pupil intentions and aspirations. A mere ten pages are devoted to a simplistic description of "recent developments affecting schooling". This section of the book attempts to cover the raising of the school-leaving age, the comprehensive re-organisation of secondary education, curriculum changes, careers guidance, occupational choice and the transition from school to work, the relationship between post-school intentions and destinations, equal opportunities legislation and so on. All in ten pages! Where reference is made to specific educational developments or earlier research it is in the briefest possible terms, and it seems inconceivable that in a book with this subject content perhaps the most significant research work undertaken is not referred to at all. This omission cannot be because earlier research has little relevance to the Scottish situation—where other work is quoted by Ryrie *et al.* little of it is exclusively Scottish. In addition, much of the most important work on subject choice and pupil intentions has emanated from Andrew McPherson and his associates at Edinburgh University, and is based on their own surveys of pupils in Scotland. This work does not get a single mention in the whole book. Similarly the failure by Ryrie *et al.* to refer either to Celia Phillips' work on subject choice at school, or to Gareth Williams' research at Lancaster University on pupil intentions or to the work of the National Children's Bureau are grave omissions. There are many others. It is to be hoped that much more attention will be given to putting the findings in context in any future publications based on this research project.

The report itself is particularly good at integrating a discussion of the grouped data obtained from the pupil interviews with verbatim comments from both pupils and teachers. In using interview quotations in this way the authors are able to illustrate effectively some of the descriptive statistics used. Most of the statistical tables in the book (frequencies and cross-tabulations only) keep the eight school groups separate. At times this is unfortunate since disaggregation of this kind does give rather small cell sizes at times (no significance tests are used.) The benefits gained from keeping the pupil data almost exclusively in eight school groups instead of two area groups are unclear. Especially so since little detailed comparative work between schools is undertaken. In addition, the representativeness of those pupils surveyed is uncertain.

In all, then, a competent, but by no means exceptional, reporting of research findings in an area that has been examined on a number of previous occasions. The report suffers greatly from the authors' failure to appreciate that a great deal of research work has already been undertaken on both subject choice and pupil intentions. In addition, the research results themselves could have been utilised more effectively. The overall impression given is that this is an interim research report that has been rushed into print. A more considered approach to any future publications emanating from the project is called for.

ALAN GORDON.

WRIGHT, H. J., and PAYNE, A. N. (1979). *Evaluation of a School Psychological Service—The Portsmouth Pattern*. Winchester: Hampshire Education Department, pp. i+69. n.p.

Educational psychologists are often tarred with the brush of expectation that theirs is a 'casualty-only' service. The School Psychological Service has been regarded and used as the 'ambulance' at the foot of life's precipices. Psychologists like Jack Wright have a long-standing ambition to find and tame those well-known beasts which topple their clients over the edge.

Teachers often say that psychologists' reports only tell them what they already know about a pupil in question (50 per cent of teachers sampled here found reports useful). It will be said that the data and arguments of this evaluation have been familiar to educational psychologists for years. Nevertheless, here is substance for wider discussion and, it is to be hoped, action.

The implicit catchwords of this report and its notion of a psychologists' working relationship with schools are 'service-delivery'. It sets a much-needed precedent for ways in which other psychological services could effectively represent themselves and the needs of their clients ('1st order clients'=children, '2nd order clients'=all adults working with them) to LEA decision-makers. It is not intended for drawing conclusions about the roles and activities of all EPs and all psychological services. The report makes it usefully clear that an SPS evaluation is particular to the demography of the region served.

The report identifies its own shortcomings and declares (section 8.1) its methodological pitfalls, with regard to the criteria which are set for effective evaluation, per se, of a service delivery programme. As a harassed EP I am distracted from making such criticisms because of the identifications with and sympathy for the aims and conclusions of this report. Equally, such criticisms would seem unrealistic when the resources (one research officer) and the time (three years) given to the project are considered.

The facts, trends and opinions reported from Portsmouth won't be novel to other EPs. However, this report provides an understanding of the role and work of psychological services essential to the modification of the misunderstandings and anachronistic stereotypings still held by their employers and users (essential reading for education administrators, education committee members and all those teaching teachers to teach).

A recurrent theme of the report is the extent to which the teachers sampled regarded and valued the psychologists' expertise with individual children and understanding of child development. The 'second-order' clients wanted this expertise applied to work with parents.

regular school visits, provision of programmes of individualised learning, more follow-up, school-based in-service training, . . . In fact, where the project gave opportunity for teachers to rank their expectations of psychologists, it appeared that they expected the resourcefulness of an educational 'Bat Man'. Furthermore, they wanted this renaissance-like acumen applied to the establishment of pupil conformity, parent-meeting (in secondary schools), development of screening techniques and provision of prescriptive learning programmes (in first and middle schools). The report notes that only one of the head teachers sampled 'expected the EP to work miracles!'. At least sixteen areas of expertise were listed by head teachers as expected features of an EP's activities and specialist repertoire. Each area has a direct bearing on a child's adaptive passage through school and each would be welcomed in any profession serving schools. The report indicates that no particular expertise would be unique to an EP and most would appear in an account of a 'good teacher'. If the report does identify unique provinces of EP expertise then they are: insightful ability to work with and for the individual child, knowledge of child development and, of course, psychometrics.

The report finds that the staffing required to fulfil the expectations held of EPs by schools would be one EP to 5,000 school-aged children. To fulfil these duties and those suggested by Warnock would require one EP to 5,000 0-year to 19-year-aged children.

Cut-backs in local government expenditure have shackled any immediate prospects that LEAs might strive to establish psychological service staffing ratios as required by the needs and expectations identified here. However, documents such as this could help to unshackle imagination and improve understanding when LEAs regard, plan the use of and staff their psychological services. It is expected that the work done in Portsmouth will be amongst the blueprints for the kind of school psychological self-determination that is necessary to counteract the threatened imposition of increased demand upon an already diluted service. Or perhaps, one of the less than a thousand EPs serving schools in England and Wales will do for school psychological services what Herriot has done for the veterinarians.

J. F. PAIN.

YOUNGMAN, M. B. (1979). *Analysing Social and Educational Research Data*. London: McGraw-Hill, pp. v+235, p. £6.95.

The last 10-15 years have seen the increasing availability of computers and packages of programmes which allow the most complicated statistical calculations to be performed at the press of a button on a teletype or visual display unit. In fact it would be surprising if there was now a single university department of education, psychology or social science which did not provide students with ready access to such facilities. Thus research workers are now spared the tedium of spending hours calculating sums of squares and cross-products etc. with a desk or hand calculator.

Consequently it is timely to question the relevance of the traditional practice of teaching the arithmetical details for carrying out statistical analysis, i.e. the approach adopted in most of the existing standard textbooks on the methodology of social and educational research. Indeed it could be argued that nowadays the most relevant textbook is one which deals only with the conceptual issues underlying statistical procedures and with the practical business of using computers to perform the different methods of analysis.

Youngman's book is one of the first, of what may turn out to be a new generation of books, written along these lines. After a preliminary chapter which provides an overview of the remainder of the book, Youngman immediately gets down to the nuts and bolts of preparing data for the computer and the details of using a computer, e.g. writing FORTRAN format statements. Subsequently all the statistical techniques presented are described in relation to the use of readily available packages of programmes such as SPSS and the author's own PMMD system.

The conceptual treatment of statistical procedures is presented over eight chapters in 150 pages or so. These chapters deal with ways of summarising the properties of distributions, measures of association and correlation etc., analysis of variance, factor analysis, multiple regression, cluster analysis, multiple discriminant function analysis, analysis of

questionnaires, construction of scales and tests, and the measurement of change. Obviously in such a small space the treatment of so many topics has to be somewhat cursory and at quite a high level. In the preface the author acknowledges that traditional texts should be used in conjunction with his own book. He also implies that a "grounding in statistics" is necessary to achieve sufficient understanding. This is my view too.

The text discusses many instructive examples of the use of the different techniques with real data derived from recent research. It also refers the reader frequently to up-to-date sources in the English and American literature for further details and discussion. There is no doubt that a great deal of thought and careful planning have gone into the preparation and writing of this book.

In conclusion the book will be welcomed by teachers of statistics in providing a good source of material to be used in combination with courses in computing. It will certainly provide some interesting new insights into the ideas underlying well-known techniques. It will be of greatest use to postgraduate students towards the end of their courses.

G. M. SEDDON.

PUBLICATIONS RECEIVED

Books for review have been selected from the following
(‘n.p.’ means ‘no price mentioned’).

- ADELSON, J. (Ed.). (1980). *Handbook of Adolescent Psychology*. Chichester: John Wiley and Sons, pp. iv+624, c. £24.50.
- AINLEY, J., and FORDHAM, A. (1979). *Between School and Adult Life: An Approach to the Evaluation of Pre-Vocational Education in TAFE*. Victoria: ACER, pp. v+207, p. \$10.00.
- ALEXOPOULOS, D. (1980). *Revision and Standardization of the Wechsler Intelligence Scale for Children—Revised (WISC-R) For the Age Range 13-15 Years in Greece*. Athens, pp. ii+364, n.p.
- AUSUBEL, O. P., SULLIVAN, E. V., and IVES, S. W. (1980). *Theory and Problems of Child Development*, (3rd Ed.). London: Academic Press, pp. iii+652, c. £16.00.
- Basic Number Diagnostic Test*. (1980). Sevenoaks: Hodder and Stoughton. Specimen set: £1.00, test: £3.00 per 20 copies, manual: 85p.
- BIGGE, M. L., and HUNT, M. P. (1980). *Psychological Foundations of Education*. (3rd Ed.). London: Harper and Row, pp. iii+545, p. £10.50.
- BOWLBY, J. (1980). *Loss: Sadness and Depression*. Vol. III of *Attachment and Loss*. London: The Hogarth Press, pp. vii+472, c. £12.00.
- BRADLEY, A., and JAMIESON, K. (1980). *Dandelion Clocks—Stories of Childhood*. London: Penguin Books, pp. 2+217, p. £1.25.
- BRENNAN, W. K. (1979). *Curricular Needs of Slow Learners*. London: Evans/Methuen Educational, pp. 6+174, p. £6.10.
- BRITISH UNIVERSITIES FILM COUNCIL (1980). *Higher Education Film Library*. London: B.U.F.C., pp. iii+56, p. £1.50.
- BRUNER, J. (1980). *Under Five in Britain*. London: Grant McIntyre, pp. viii+228, c. £9.95, p. £3.95.
- BRYANT, B., HARRIS, M., and NEWTON, D. (1980). *Children and Minders*. London: Grant McIntyre, pp. vii+244, c. £9.95, p. £3.95.
- CARR, J. (1980). *Helping Your Handicapped Child—A Step-by-Step Guide to Everyday Problems*. London: Penguin Books, pp. 10+271, p. £1.95.
- CHAZAN, M., LAING, A. F., SHACKLETON, B. M., and JONES, G. (1980). *Some of our Children—The Early Education of Children with Special Needs*. London: Open Books, pp. vii+260, p. £5.95.
- CLAESON, B. H. (1980). *A Guide to Sex for Young People: Boy Girl Man Woman*. London: Penguin Books, pp. 9+142, p. £1.25.
- CLAXTON, G. (Ed.). (1980). *Cognitive Psychology—New Directions*. London: Routledge and Kegan Paul, pp. 1+338, c. £12.50.
- CLIFT, P., CLEAVE, S., and GRIFFIN, M. (1980). *The Aims, Role and Deployment of Staff in the Nursery*. Windsor: NFER, pp. 8+224, p. £6.75.
- COOPER, D. (1980). *The Language of Madness*. London: Penguin Books, pp. 12+182, p. £1.50.
- COUNCIL OF EUROPE (1980). *Information Bulletin on Legal Activities—within the Council of Europe and in Member States*. Strasbourg: Council of Europe, pp. 1+60, n.p.
- CRUICKSHANK, W. M. (Ed.). (1980). *Psychology of Exceptional Children and Youth*. (4th Ed.). Hemel Hempstead: Prentice/Hall International, pp. v+595, c. £11.65.
- DICAPRIO, N. S. (1980). *Adjustment—Fulfilling Human Potentials*. Hemel Hempstead: Prentice/Hall International, pp. iii+501, p. £8.40.
- DODWELL, P. C. (Ed.). (1980). *New Horizons in Psychology*. (2nd Ed.). London: Penguin Books, pp. 10+323, p. £2.75.
- DONAHOE, J. W., and WESSELLS, M. G. (1980). *Learning, Language and Memory*. London: Harper and Row, pp. viii+513, c. £11.00.
- EGAN, K. (1980). *Educational Development*. Oxford: Oxford University Press, pp. 4+174, p. £2.75.
- FRANCKE, L. B. (1980). *The Ambivalence of Abortion*. London: Penguin Books, pp. 3+261, p. £1.95.
- GALTON, M., SIMON, B., and CROLL, P. (1980). *Inside the Primary Classroom*. Henley-on-Thames: Routledge and Kegan Paul, pp. v+203, c. £8.95, p. £4.95.
- GARLAND, C., and WHITE, S. (1980). *Children and Day Nurseries*. London: Grant McIntyre, pp. vii+128, c. £7.95, p. £2.95.

- GOODWIN, W. L., and DRISCOLL, L. A. (1980). *Handbook for Measurement and Evaluation in Early Childhood Education*. London: Jossey-Bass pp. vii+632, c. £22.00.
- GRITTNER, F. M. (Ed.) (1980). *Learning A Second Language*. Chicago: University of Chicago Press, pp. iv+241, c. £7.50.
- HARBISON, J., and HARBISON, J. (1980). *A Society Under Stress—Children and Young People in Northern Ireland*. London: Open Books, pp. vii+200, c. £9.95.
- HERSOV, L. A., and BERGER, M. (Ed.) (1980). *Language and Language Disorders in Childhood*. Oxford: Pergamon Press, pp. v+183, p. \$9.00.
- HILL, W. F. (1980). *Learning—A Survey of Psychological Interpretations*. Andover: Methuen, pp. vii+291, p. £3.95.
- HJELLE, L. A., and ZIEGLER, D. J. (1980). *Personality—Theories, Basic Assumptions, Research, and Applications*. Maidenhead: McGraw Hill, pp. viii+362, p. £4.95.
- HOLMES McDUGALL Seminars. (1979). *Teaching Reading*. Edinburgh: Holmes McDougall, set of booklets p. £75.00.
- HOWE, M. J. A. (1980). *The Psychology of Human Learning*. London: Harper and Row, pp. vii+448, c. £10.50.
- HUGHES, M. et al. (1980). *Nurseries Now. A Fair Deal for Parents and Children*. London: Penguin Books, pp. 8+284, p. £1.95.
- JAFFE, J., PETERSEN, R., and HODGSON, R. (1980). *Addictions—Issues and Answers. A Life Cycle Book*. London: Harper and Row, pp. 4+128, c. £4.95, p. £2.25.
- JENSEN, A. R. (1980). *Bias in Mental Testing*. London: Methuen, pp. vii+786, c. £15.00.
- JOHNSON, M. (Ed.) (1980). *Towards Adolescence—The Middle School Years*. Chicago: University of Chicago Press, pp. iv+338, c. £9.00.
- KARMEL, P. (1979). *Youth, Education and Employment*. Victoria: ACER, pp. 3+19, p. \$2.00.
- KLAUSMEIER, H. J. and ASSOCIATES. (1980). *Cognitive Learning and Development: Piagetian and Information-Processing Perspectives*. London: Harper and Row, pp. v+370, c. £16.25.
- LAWTON, D. (1980). *The Politics of the School Curriculum*. Henley-on-Thames: Routledge and Kegan Paul, pp. vii+152, c. £6.95, p. £3.95.
- LAYTON, J. R. (1980). *The Psychology of Learning to Read*. London: Academic Press, pp. vii+356, c. £12.20.
- LERNER, R. M., and SPANER, G. B. (1980). *Adolescent Development—A Life-Span Perspective*. Maidenhead: McGraw Hill, pp. vii+537, c. £10.15.
- MAHONEY, M. J. (1980). *Abnormal Psychology—Perspectives on Human Variance*. London: Harper and Row, pp. v+590, c. £12.50.
- MAIN, A. (1980). *Encouraging Effective Learning*. Edinburgh: Scottish Academic Press, pp. 2+76, p. £3.25.
- MUSSEN, P. H., CONGER, J. J., and KAGAN, J. (1980). *Essentials of Child Development and Personality*. London: Harper and Row, pp. ix+451, c. £10.50.
- MCGHEE, P. E., and CHAPMAN, A. J. (Eds.) (1980). *Children's Humour*. Chichester: John Wiley and Sons, pp. v+321, c. £14.00.
- NATIONAL TECHNICAL INSTITUTE FOR THE DEAF (1980). *Partnership in Learning*. Rochester: N.T.I.D., pp. 1+28, n.p.
- NORTON, M. (1980). *Bedknob and Broomstick*. London: Penguin Books, pp. 10+188, p. £0.85p.
- PETRIE, C. (1980). *The Nowhere Boys—A Comparative Study of Open and Closed Residential Placement*. Farnborough: Saxon House, pp. v+163, c. £8.50.
- PUFF, C. R. (Ed.) (1980). *Memory Organisation and Structure*. London: Academic Press, pp. v+411, c. £15.40.
- RAPOPORT, R. R. (1980). *Growing Through Life—A Life Cycle Book*, London: Harper and Row, pp. 4+128, c. £4.95, p. £2.25.
- RESCORLA, R. A. (1980). *Pavlovian Second-Order Conditioning—Studies in Associative Learning*. Chichester: John Wiley and Sons, pp. v+120, c. £8.50.
- REST, J. R. (1979). *Development in Judging Moral Issues*. Minneapolis: University of Minnesota Press, pp. vii+305, c. \$19.50.
- RICHARDS, M. (1980). *Infancy—World of the Newborn. A Life Cycle Book*. London: Harper and Row, pp. 4+128, c. £4.95, p. £2.25.
- ROBERTS, T. B. (1980). *Consciousness, Psychology, and Education—A Speculative Study*. Illinois: New Learning Publishers, pp. 2+38, p. £4.00.

- RODERICK, G. and STEPHENS, M. (Ed.) (1980). *Higher Education For All?* Lewes, Sussex: Falmer Press, pp. 9+206, c. £9.85.
- SALMON, P. (Ed.) (1980). *Coming to Know*. Henley-on-Thames: Routledge and Kegan Paul, pp. v+188, p. £5.95.
- SHEARER, A. (1980). *Handicapped Children in Residential Care—A Study of Policy Failure*. London: Bedford Square Press, pp. 1+114, p. £4.95.
- SOLSO, R. L. (1960). *Cognitive Psychology*. London: Harcourt Brace Jovanovich, pp. vi+499, c. £10.75.
- STEVENS, A. (1980). *Clever Children in Comprehensive Schools*. London: Penguin Books, pp. 7+171, p. £1.25.
- STIERER, B. (1979). *Reading and Language Projects in Local Education Authorities. (A National Directory of LEA Literacy Research and Curriculum Development)*. Brighton: The Literacy Centre, Brighton Polytechnic, pp. 5+129, p. £3.00.
- STOTT, D. H. (1980). *Delinquency and Human Nature*. (2nd Ed.). Sevenoaks: Hodder and Stoughton, pp. v+432, c. £9.95.
- SUTTON-SMITH, B. (Ed.) (1980). *Play and Learning*. Chichester: John Wiley and Sons, pp. vii+335, c. £14.60.
- SYLVA, K., ROY, C., and PAINTER, M. (1980). *Childwatching at Playgroup and Nursery School*. London: Grant McIntyre, pp. viii+270, c. £9.95 and p. £4.25.
- TANZ, C. (1980). *Cambridge Studies in Linguistics. 26. Studies in the Acquisition of Deictic Terms*. Cambridge: Cambridge University Press, pp. vi+184, c. £15.00.
- TAYLOR, P. H. (Ed.) (1980). *New Directions in Curriculum Studies*. Lewes, Sussex: Falmer Press, pp. v+180, c. £7.95.
- TROWN, A., and NEEDHAM, G. (1980). *Reduction in Part-Time Teaching: Implications for Schools and Women Teachers*. Manchester: Equal Opportunities Commission, pp. 1+130, n.p.
- TUMA, D. T., and REIF, F. (Eds.) (1980). *Problem Solving and Education: Issues in Teaching and Research*. Chichester: John Wiley and Sons, pp. iii+212, c. £10.00.
- UNDERWOOD, G., and STEVENS, R. (Eds.). (1979). *Aspects of Consciousness*. Vol. 1. London: Academic Press, pp. v+251, c. £12.40.
- TH. VAN DER KAMP, K. J. LANGERAK, W. F., and GRUIJTER, D. N. M. (1980). *Psychometrics for Educational Debates*. Chichester: John Wiley and Sons, pp. vi+337, c. £18.40.
- WALKER, D. A. (1980). *Understanding Pictures: A Study in the Design of Appropriate Visual Materials for Education in Developing Countries*. Univ. of Massachusetts: Center for International Education, pp. iii+380, p. \$6.50.
- WARREN, M. A. (1980). *The Nature of Woman. An Encyclopedia and Guide to the Literature*. Point Reyes, Calif.: Edgepress, pp. i+696, c. \$20.00.
- WATSON, P. (1980). *War on The Mind—The Military Uses and Abuses of Psychology*. London: Penguin Books, pp. 7+414, p. £2.95.
- WHITE, S. and NOTKIN, B. (1980). *Childhood—Pathways of Discovery. A Life Cycle Book*. London: Harper and Row, pp. 4+127, c. £4.95, p. £2.25.
- WINNICOTT, D. W. (1980). *Playing and Reality*. London: Penguin Books, pp. xii+194, p. £1.95.
- WINNICOTT, D. W. (1980). *The Piggle—An Account of the Psychoanalytic Treatment of a Little Girl*. London: Penguin Books, pp. vii+201, p. £1.95.
- WOLFENDALE, S. and BRYANS, T. (1980). *Identification of Learning Difficulties: A Model for Intervention*. Stafford: N.A.R.E., pp. 1+312, n.p.
- WOLKIND, S. (Ed.) (1979). *Medical Aspects of Adoption and Foster Care*. London: S.I.M.P., pp. v+102, c. £6.50.

AUTHOR INDEX TO VOLUME 50, 1980

ADAMS, M.: See Nisbet	
ARTHUR, J.: See Nisbet	
BALDWIN, R.: See McKeachie	
BAR-TAL, D.: The Relationship between Locus of Control and Academic Achievement, Anxiety and Level of Aspiration	53
BAR-TAL, D.: See Raviv	
BAR-TAL, Y.: See Raviv	
BAR-ZOHAR, Y.: See Bar-Tal	
BEDWELL, C. H.: Visual and Ocular Control Anomalies in Relation to Reading Difficulty ...	61
BOYES, M. C.: See Siegal	
BRADSHAW, J.: See Silva	
CARTER, M.: See Francis	
CATTELL, R. B.: The Heritability of Fluid (g_f) and Crystallised (g_c) Intelligence estimated by a Least Squares Use of the MAVA Method	253
CHEN, M.: See Bar-Tal	
COE, K. J.: See Fry (both entries)	
DAUGHERTY, M.: See McKeachie	
DOWLING, J. R.: Adjustment from Primary to Secondary School: a One Year Follow-up ...	26
EDWARDS, P.: See Harvey	
ELLERMAN, D. A.: Self-Regard of Primary School Children: some Australian Data ...	114
EVANS, H.: See Fontana	
EVANS, L.: WISC Performance Scale and Coloured Progressive Matrices with Deaf Children	216
FESHBACH, N. D.: See Phinney	
FONTANA, D.: Mode of Stimulus Presentation and Short-Term Memory Efficiency in Primary School Children	229
FRANCIS, L.: Church Aided Secondary Schools, Religious Education as an Examination Subject and Pupil Attitude towards Religion	297
FRY, P. S.: Achievement Performance of Internally and Externally Oriented Black and White High School Students under Conditions of Competition and Co-operation Expectancies	162
FRY, P. S.: Interaction amongst Dimensions of Academic Motivation and Classroom Social Climate: a Study of the Perceptions of Junior High and High School Pupils	33
FUELOEP, S.: See Lawson	
GARDNER, P. L.: A Transmission-Interpretation Scale	186
GRANT, R.: See Bedwell	
HARTLEY, J.: The Effect of Headings on Children's Recall from Prose Text	304
HARVEY, T. J.: Children's Expectations and Realisations of Science	74
HEWISON, J.: Parental Involvement and Reading Attainment	209
HICKS, C.: The ITPA Visual Sequential Memory Task: an Alternative Interpretation and the Implications for Good and Poor Readers	16
HORNSBY, B.: The Effects of a Dyslexia-centred Teaching Programme	236
IVES, S. W.: The Use of Orientations in Children's Drawings of Familiar Objects: Principles versus Percepts	295
KENELY, J.: See Hartley	
KFIR, D.: See Bar-Tal	
LAMBERT, W. B.: See Russell	
LAWSON, M. J.: Understanding the Purpose of Strategy Training	175
LIN, Y-G.: See McKeachie	
LINDSAY, G. A.: The Infant Rating Scale	97
MACKAY, C. K.: See Millar	
MCDERMOTT, P. A.: Principal Components Analysis of the Revised Bristol Social Adjustment Guides	223
McKEACHIE, W. J.: Using Student Ratings and Consultation to Improve Instruction ...	168



McKEOWN, J. R.:	See Bedwell	
MAQSUD, M.:	Effects of Personal Lecture Notes and Teacher-Notes on Recall of University Students	289
MAQSUD, M.:	Extraversion, Neuroticism, Intelligence and Academic Achievement in Northern Nigeria	71
MERRY, R.:	The Keyword Method and Children's Vocabulary Learning in the Classroom...	123
MILES, T. R.:	See Hornsby	
MILLAR, C.:	Conflict between Cues in Number Conservation Tasks	188
MOFFETT, M. M.:	See McKeachie	
NEIGLER, C.:	See McKeachie	
NISBET, J.:	A Simple Form of Tailored Testing	301
NORK, J.:	See McKeachie	
OWEN, G.:	See Hartley	
OWENS, L.:	The Development of a Co-operative, Competitive, and Individualised Learning Preference Scale for Students	147
PHINNEY, J. S.:	Non-Directive and Intrusive Teaching Styles of Middle- and Working-class English Mothers	2
RAVIV, A.:	Causal Perceptions of Success and Failure by Advantaged, Integrated and Disadvantaged Pupils	137
RAVIV, Alona:	See Raviv	
ROGERS, C. G.:	The Development of Sex Differences in Evaluations of Others' Successes and Failures	243
RUSSELL, G. W.:	Sherlock Holmes meets Othello: a MDS Analysis of Literary Characters	277
SIEGAL, M.:	The Comparative Effects of Adult and Peer Group Influence on Children's Moral Judgments... ..	105
SILVA, P. A.:	Some Factors Contributing to Intelligence at Age of School Entry	10
STRATON, R. G.:	See Owens	
TAYLOR, S. M.:	See Gardner	
TIZARD, J.:	See Hewison	
TRUEMAN, M.:	See Hartley	
WALZ, M.:	See McKeachie	
WEBB, N. M.:	An Analysis of Group Interaction and Mathematical Errors in Heterogeneous Ability Groups	266
YOUNG, G. M.:	The Attitudes of Teachers in Further Education	181
YOUNGMAN, M. B.:	Some Determinants of Early Secondary School Performance	43

NOTES TO CONTRIBUTORS

Manuscripts are accepted on the understanding that they report unpublished work that is not under consideration for publication elsewhere, and that if accepted for the *Journal*, the work will not be published again in the same form, in any language, without the consent of the Editor.

The usual range of length of papers is from 2,500 to 4,000 words. Papers should be in a clear style and free from jargon. Statistical matter (formulae, tables, discussion of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as mean, percentiles, standard deviations, correlation, chi-squared, standard error, critical ratio, significance and reliability, regression and factor analysis, can be employed in the text and need not be explained.

Short research notes will also be published, reporting work of two kinds: (i) results of sufficient moment to merit publication in advance of a more comprehensive paper; (ii) work which substantially confirms or extends existing knowledge but which does not justify an extensive paper.

Contributors can help to avoid delays in publication by observing certain guiding principles:

1—Papers should be written with the utmost conciseness consistent with clarity. If the editor or a referee takes the wrong meaning from or fails to understand a passage, the fault is the author's. Unless the editor can grasp the meaning of a sentence unequivocally, it cannot be assumed that other readers can.

2—A paper should be written only when a piece of work is rounded off. A report of a pilot study, or a series of papers on the same subject written as the results come to hand, is unlikely to be accepted for publication. A comprehensive paper, particularly if it includes reports of several experiments or analyses, or replications on further samples, which give consistent results, will be given favourable consideration.

3—Experimental reports should generally follow the standard pattern of presentation, e.g.:

- (a) A short summary at the start.
- (b) A statement of aims or objectives and a review of relevant previous work as an introduction to the topic—conciseness and strict selection of relevant aspects are particularly important here.
- (c) A clear statement of the sample, the procedure, the tests, the experimental material, in sufficient detail to permit a reader to replicate the study if he so wishes.
- (d) A factual statement of results, without speculative or interpretative comment or discussion.
- (e) Either a brief list of general conclusions, not repeating results already listed, or a discussion, which should be strictly limited to discussing the results (if this is necessary) and which should not involve mere recapitulation or speculation beyond the scope of the inquiry.
- (f) Acknowledgments.
- (g) References.

4—The summary (which is often copied by abstracting agencies) should be intelligible without reference to the paper. It should give an indication of the aim, scope and results of the inquiry, preferably in about three sentences. It should mention the size and nature of the sample, and should not include vague phrases such as "The results were analysed . . ." or "The implications of the findings are discussed . . ." Even for a full-length paper, the summary should not exceed 200 words.

5—Figures and tables should be selected to illustrate or summarise points made in the text. The same data should not be reported both in table and figure; tables are generally preferable because they are much more economical. Tables should contain selected data: only the most important quantitative results can be reported in full.

6—Authors should consult a current issue of the *Journal* in order to follow the conventions applied in the *Journal*—the use of headings, numbering of tables, the absence of footnotes, the impersonal style of reporting, omission of dots in standard abbreviations, reference to subjects as children and not as Ss, the lay-out of references and so on. The British Psychological Society pamphlet, *Suggestions to Authors*, should be consulted. References should be cited by author and date in the text, and arranged in alphabetical order in the list of references at the end. All measurements should be stated in SI units.

7—A manuscript should be headed with the title of the paper, the names of the authors (female authors should put one of their given names in full, male authors initials and surname) and, in brackets, a short address identifying the institution where the work was done. Titles should be short. If the title exceeds 45 letters and spaces in length, a shortened version, suitable as a running title, should be added.

8—The manuscript should be typed in double-spacing with wide margins on one side of the paper which is suitable for ink corrections and printer's instructions—not duplicated or on flimsy paper. Drawings should be in Indian ink on cartridge paper. Manuscripts which require re-typing before they can be handled by the printer have to be sent back to the authors. The *Journal* has no paid staff.

Authors receive 50 copies of their papers free (delivered about six weeks after publication). Extra copies may be had at cost price if the order is given when the proof is returned.

The *Journal* will adopt the International System of Units (SI) based on the units: metre, kilogramme, second, ampere, Kelvin and candela. Authors who refer to physical measurements in their papers should now use SI units except where it would be absurd to do so. Further information about SI units, including conversion tables, is contained in the revised *Suggestions to Authors* pamphlet issued by the British Psychological Society and obtainable, price 25p (USA \$1.00) post free, from the Cambridge University Press.

The quarterly *Educational Psychology* will provide an international forum for the discussion and rapid dissemination of research findings in psychology relevant to education.

The journal will place particular emphasis on the publishing of papers reporting applied research based on experimental and behavioural studies. It is hoped that the journal will quickly become a primary source for papers ranging from pre-school to tertiary provision, concentrating primarily on 'normal' children in 'normal' educational settings. The prompt publication of high-quality articles will be the journal's first priority. All contributions will be submitted 'blind' to at least two independent referees before acceptance for publication.

Contributions are invited immediately (normally no more than 8000 words, in three copies please).

The Joint Editors,
Educational Psychology,
Department of Educational
Psychology, University of
Birmingham, PO Box 363,
Birmingham B15 2TT, England.

The journal will be published quarterly, in March, June, September and December. Vol 1 No. 1 March 1981. ISSN 0144-3410

Joint Editors

Dr **Kevin Wheldall** &
Dr **Richard Riding**
University of Birmingham

Editorial Board

Dr **Wesley C. Becker**, University of Oregon
Dr **Ann Clarke**, University of Hull
Professor **Margaret Clark**, University of Birmingham
Dr **Cynthia P. Deutsch**, New York University
Dr **Margaret Donaldson**, University of Edinburgh
Dr **David Fontana**, University of Cardiff
Professor **Ted Glynn**, University of Auckland
Professor **Reg Marsh**, University of Wellington
Professor **Peter Mittler**, University of Manchester
Dr **Sohan Modgil**, Brighton Polytechnic
Professor **Walt Musgrove**, University of South Florida
Professor **James Ward**, Macquarie University

Statistical Editor Robert Lambourne
Sub-Editor Hazel Benner

Subscription rate one year £30 00 (\$85 00), post free. Individuals may subscribe at a 50% discount—please write to publishers for details of Personal Subscription Plan. Carfax Publishing Company, Haddon House, Dorchester-on-Thames, Oxford OX9 8JZ, England

EDUCATIONAL PSYCHOLOGY

An international journal of experimental educational psychology



The British Journal of Psychology

Volume 71, Part 3, August 1980

Contents

- S. F. Walker.** Lateralization of functions in the vertebrate brain: A review
G. H. Bower & G. Clark-Meyers. Memory for scripts with organized vs. randomized presentations
P. H. K. Seymour. Semantic and structural coding of the months
P. Nalsh. Phonological recoding and the Stroop effect
N. Kapur. Recognition of word associates in semantic paralexia
K. Millar, B. G. Styles & D. G. Wastell. Time of day and retrieval from long-term memory
G. J. Walker-Smith. Memorizing facial identity, expression and orientation
G. Jahoda. Sex and ethnic differences on a spatial—perceptual task: Some hypotheses tested
A. W. Boykin & J. Harackiewicz. Epistemic curiosity and incidental recognition in relation to degree of uncertainty
A. Koriat & R. Melkman. Individual differences in memory organization as related to word-association, object-sorting and word-matching styles
Book reviews

Volume 71 (1980) £35.00 (US\$77.00). Single parts cost £11.00 (US\$24.00)

Volume 70 (1979) £26.00

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK

The British Journal of Medical Psychology

Volume 53, Part 3, September 1980

Contents

- Comments on Watts (1980) 'Clinical judgement and clinical training': An editorial contribution
D. WIII. Psychoanalysis as a human science
R. Littlewood. Anthropology and psychiatry—An alternative approach
L. Munoz. Exile as bereavement: Socio-psychological manifestations of Chilean exiles in Great Britain
D. Whitwell. The use of psychotherapy with in-patients
R. Harris & M. W. Linn. Differential response of the need to be held
S. J. Biggs, B. Rosen & A. B. Summerfield. Video-feedback and personal attribution in anorexic, depressed and normal viewers
D. R. Schwoon & M. C. Angermeyer. Congruence of personality assessments within families with a schizophrenic son
D. Fontana. Some standardization data for the Sandler—Hazan Obsessionality Inventory
G. A. Fava, G. I. Perini, P. Santonastaso & C. Veller Fornasa. Life events and psychological distress in dermatologic disorders
Book reviews
Other publications received

Volume 53 (1980) £28.00 (US\$62.00). Single parts cost £9.00 (US\$20.00)

Volume 52 (1979) £21.00 (US\$45.00)

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK

Journal of Occupational Psychology

An international journal of research into people at work. Published quarterly, covering industrial, organizational, engineering, vocational and personal psychology, as well as behavioural aspects of industrial relations and human factors. Innovative or interdisciplinary approaches with a psychological emphasis are particularly welcome.

Contents of Volume 53, Part 2, June 1980

A. C. West Introducing participation: An example from the British ports industry

A. T. Welford On the nature of higher-order skills

C. Lewis Investigating the employment interview. A consideration of counselling skills

D. E. Terpstra & J. M. Larsen, Jr A note on job type and applicant race as determinants of hiring decisions

G. E. O'Brien & P. Dowling The effects of congruency between perceived and desired job attributes upon job satisfaction

I. T. Robertson & R. M. Mindel A study of trainability testing

S. J. Hapworth Moderating factors of the psychological impact of unemployment

J. F. Hartley The impact of unemployment upon the self-esteem of managers

Occupational psychology through autobiography: **John Parry**

Book reviews

Volume 53 (1980) £25.00 (US\$55.00) Single parts cost £8.00 (US\$17.00)

Orders to:

The British Psychological Society

The Distribution Centre, Blackhorse Road, Letchworth, Hertfordshire SG6 1HN, UK

RICHARD LANSDOWN

More than Sympathy

The everyday needs of sick and handicapped children and their families

This book is concerned with the proportion of young children whose stay in hospital is lengthy and whose difficulties are likely to be more or less life long. Written for those who are training to work with these children and their families, it takes as a starting point the notion that many of the problems that arise can be predicted and that to know something of the child's condition and how it may affect behaviour is to be part way towards coping.

224 pages

Hardback 0 422 76630 5 £7.95

Paperback 0 422 76640 2 £3.50



TAVISTOCK

METHUEN



JOHANNA TURNER

Made for Life

Coping, competence and cognition

Made for Life considers the relationship between feeling able to cope and being able to learn — that is, the interdependence of affect and cognition in children under five. It argues that in order to make full use of his cognitive capacities the child must first develop the belief that he is able to cope and be effective.

192 pages

Hardback 0 416 72690 9 £8.00

Paperback 0 416 72700 X £3.95

ENCOURAGING EFFECTIVE LEARNING

By **ALEX MAIN**

There are many books on study methods, but they are not as widely consulted as might be supposed. Students tend to turn to their teachers for help with study difficulties. Yet, many teachers, tutors and lecturers feel limited in the extent of the advice they can give. So far there has been no book available which deals with this aspect of a teacher's work.

Now, teachers in secondary, further and higher education can turn to Alex Main's book, *Encouraging Effective Learning*, for help in this difficult area.

Alex Main has developed a successful study counselling service at the University of Strathclyde, and his work has attracted wide interest in this country and abroad. His book describes how other teachers can use his approach in advising and counselling students who have study difficulties.

This book describes techniques and activities which require no specialised training. It creates an understanding of the styles of learning which different students adopt, and the difficulties they can face: and it shows how the sympathetic teacher can do much to relieve their anxiety and improve their methods.

The book also contains a guide to books, pamphlets and audio-visual materials which provide additional help to students and useful resources for the teacher.

SBN 7073 0268 4

76pp

paperback £3.25



SCOTTISH ACADEMIC PRESS

33 MONTGOMERY STREET, EDINBURGH

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

available from the Society's offices

St Andrews House, 48 Princess Road East, Leicester, LE1 7DR

Careers in Psychology	50p
Postgraduate Qualifications and Courses in Psychology	40p
Postgraduate Careers in Clinical Psychology	20p
Postgraduate Careers in Educational Psychology	20p
Postgraduate Careers in Occupational Psychology	20p
Postgraduate Careers in The Prison Service	20p
Postgraduate Careers in Social Psychology	20p

PRESIDENTIAL ADDRESSES

1968 Dissonance and Stress in the Teaching of Psychology: Some Reflections of an Accidental Psychologist: Professor George Seth	15p
1969 Changing Horizons: An Essay in Autobiography: Dr Boris Semeonoff	15p
1970 Choosing: Professor R. J. Audley	15p
1971 In Search of an Idiographic Psychology: Professor H. Gwynne Jones	15p
1972 Psychology Today and Tomorrow: Professor Harry Kay	15p
1973 Psychology in Society: Ends or End: Professor Max Hamilton	15p
1974 On Taking Sides: Professor B. M. Foss	15p
1975 Thought and the Brain: Professor O. L. Zangwill	15p
1976 Psychology and Social Policy: Professor Jack Tizard	15p
1977 The Scientific/Applied Debate in Psychology: A Contribution: May A. Davidson	25p
1978 Predicting Human Development: Professor A. D. B. Clarke	25p

C. S. MYERS LECTURES

1964 Remembering Dr Myers: Professor F. C. Bartlett	15p
1965 Personality and Experimental Psychology: Professor H. J. Eysenck	15p
1967 Practice at liking: A study in experimental aesthetics: Professor D. W. Harding	15p
1968 What is potential ability?: Professor P. E. Vernon	15p
1970 In defence of empirical psychology: Dr D. E. Broadbent	15p
1971 C. S. Myers in retrospect: Professor Alec Rodger	15p
1973 On applied psychology in Britain: Professor George Drew	15p
1975 Measuring psychiatric morbidity in a South Indian population: Dr G. M. Carstairs	15p
1976 Quantitative subjective assessments are almost always biased: Sometimes completely misleading: Dr E. C. Poulton	15p
1977 The neuropsychology of anxiety: Dr J. A. Gray	15p

REPORTS TO GOVERNMENT DEPARTMENTS

To the Royal Commission on the Law Relating to Mental Illness and Mental Deficiency 1954-1957	15p
To the Departmental Committee on the Adoption of Children	15p
To the Butler Committee on the Law relating to the Mentally Abnormal Offender	15p
Comments on the DHSS Review of the Mental Health Act 1959: Professional Affairs Board: <i>Bulletin</i> (1977), 30, 223-226	15p
Evidence to the Data Protection Committee (Home Office): Professional Affairs Board: <i>Bulletin</i> (1977), 30, 224-248	15p
Comments on the Report of the Local Committee on Child Health Services: Professional Affairs Board: <i>Bulletin</i> (1978), 31, 157-160	15p
Evidence to the DHSS Working Group in Primary Health Care Teams: Professional Affairs Board	25p
Comments on DHSS Discussion Document "Marriage Matters": Professional Affairs Board	25p
Comments on Interim Health Building: Note 36, Health Centres: Professional Affairs Board	25p
Comments on the Report of the Committee of Enquiry into Mental Handicap Nursing and Care (Jay Report): Professional Affairs Board	25p
Comments on the Report of the Committee of Enquiry into Normansfield Hospital: Professional Affairs Board	25p
Comments on Home Office Green Paper "Youth Custody and Supervision": Professional Affairs Board	25p

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

(continued)

REPORTS TO GOVERNMENT DEPARTMENTS—continued

Evidence to the Royal Commission on the National Health Service: Professional Affairs Board: March 1977	15p
Comments on the Draft Document on Prevention in the Child Health Services: Professional Affairs Board: May 1978	15p
Evidence to the Home Office Review Committee on the Prison Psychological Service: Professional Affairs Board: May 1978	15p
Comments on the draft health circular on Health Centres and Other Provisions for Primary Health Care: Professional Affairs Board: May 1978	15p
Comments on the DHSS Document on Health Service Complaints Procedure: Professional Affairs Board: August 1975	25p
Comments on "A Happier Old Age": A discussion document from the DHSS and Welsh Office: Professional Affairs Board: September 1978	25p
Manpower Survey: Psychologists and Mental Handicap: Professional Affairs Board: November 1978	25p
Comments on N.D.G. Pamphlet No. 5: Professional Affairs Board: November 1978	25p
Evidence to the Royal Commission on Criminal Procedure: Professional Affairs Board: November 1978	25p
Response to the Consultative Document on the Warnock Report: Professional Affairs Board: March 1979	25p
Evidence to the DHSS Working Party on Observation and Assessment: Professional Affairs Board: 1979	25p

OTHER REPORTS

Principles governing the Employment of Psychological Tests and Clinical Instruments	15p
Psychological Tests: A Statement by The British Psychological Society	15p
Technical Recommendations for Psychological and Educational Tests	15p
Teaching Psychology to Nurses (DCP Report)	15p
The Work of the Conference Programme, by Dr George Delafield	15p
The Teaching of Psychology in Schools: Joint Statement by The British Psychological Society and the Association for the Teaching of Psychology: September 1977	20p
British Psychological Education: Scientific Affairs Board: November 1977	25p
Ethical Principles for Research with Human Subjects: Scientific Affairs Board: April 1978	20p
Comments on the Proposed New CCETSW Training Programme: Professional Affairs Board: May 1978	25p
Report of the Working Party on the Sex Discrimination Act: Professional Affairs Board: May 1978	25p
Report of the Working Party on Behaviour Modification: Professional Affairs Board: September 1978	25p
Report of the Working Party on Abuses of Psychology for Political Purposes: Professional Affairs Board: October 1978	25p
Notes for guidance in planning short courses on Psychological Testing: Standing Committee on Test Standards: February 1979	25p
Postgraduate training in Clinical Psychology: Optimal length and format of training: Professional Affairs Board: <i>Bulletin</i> (1977), 30, 284-285	25p
Computer Use in British Psychology Departments: David Hale for Scientific Affairs Board: <i>Bulletin</i> (1977), 30, 410-413	25p
Repertory Grid Methods: Mathematical & Statistical Psychology Section	50p
Monitoring Financial Support for Psychological Research: Scientific Affairs Board: <i>Bulletin</i> (1978), 31, 3-5	25p
Psychological Services for Children: Professional Affairs Board: <i>Bulletin</i> (1978), 31, 125-127	20p
Report of Working Party on Animal Experimentation: Scientific Affairs Board: <i>Bulletin</i> (1979), 32, 44-52	25p
Comments to Royal College of Psychiatrists on the Role of the Community Psychiatric Nurse: Professional Affairs Board: July 1979	25p
Criteria for the Assessment of Postgraduate Training Courses in Clinical Psychology: Professional Affairs Board: November 1979	25p
Report on the Society's Working Party on <i>Corporal Punishment in Schools</i> : a Discussion Document (May 1980)	£3.00p

price includes postage

Oxford University Press

Educational Development

Kieran Egan

'Except among educational psychologists, educational psychology has a bad name . . . It is therefore an exceptional pleasure to read a book on such themes which, for once, exhilarates . . . Professor Egan's analysis and theory are highly original . . . Like all good books this one makes one think well beyond the confines of its immediate concerns.'
The Times Educational Supplement Paper covers £2.75

Educational Psychology in the Classroom

Henry Clay Lindgren

The aim of this textbook is the clear presentation of the basic principles of educational psychology and their application in the classroom.
Sixth edition illustrated paper covers £10.95

PUBLICATIONS OF THE BRITISH PSYCHOLOGICAL SOCIETY

The British Journal of Psychology
The British Journal of Medical Psychology
Journal of Occupational Psychology
The British Journal of Social and Clinical Psychology
The British Journal of Mathematical and Statistical Psychology
The British Journal of Educational Psychology

The British Journal of Psychology, *The British Journal of Medical Psychology*, *Journal of Occupational Psychology* and *The British Journal of Social and Clinical Psychology* appear in quarterly parts. The 1980 subscription prices, payable in advance, are £35.00 (US \$77.00 in the USA and Canada) for *The British Journal of Psychology*; £33.00 (US \$74.00) for *The British Journal of Social and Clinical Psychology*; £28.00 (US \$62.00) for *The British Journal of Medical Psychology* and £25.00 (US \$55.00) for *Journal of Occupational Psychology*.

The British Journal of Mathematical and Statistical Psychology is issued twice a year, in May and November. The subscription rate, payable in advance, is £25.00 per volume post free (US \$55.00 in the USA).

Subscriptions to any of these journals may be sent to a bookseller or subscription agent, or to The British Psychological Society, The Distribution Centre, Blackhorse Road, Letchworth, Herts SG6 1HN, UK. A special price of £130 (\$290) for all five journals going to the same address is offered for 1980.

The British Journal of Educational Psychology is issued three times a year, in February, June and November. The prepaid subscription is £12.00 a year, post free, or £5.00 each number. The prepaid subscription price in the USA is \$25.00 post free. Subscriptions should be sent to the Scottish Academic Press, 33 Montgomery Street, Edinburgh, EH7 5JX.

MANUSCRIPTS

For *The British Journal of Psychology* should be sent to Professor M. Coltheart, Department of Psychology, Birkbeck College, Malet Street, London WC1E 7HX.

For *The British Journal of Medical Psychology* should be sent to Professor J. P. Watson, Department of Psychiatry, Guy's Hospital Medical School, London Bridge, London, SE1.

For *Journal of Occupational Psychology* should be sent to Professor P. Warr, MRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For *The British Journal of Social and Clinical Psychology* should be sent to G. M. Stephenson (Social Psychology Editor), Professor of Social Psychology, Elliot College, University, Canterbury, Kent, CT2 7NS; or to Dr. D. Shapiro (Clinical Psychology Editor), MRC Social and Applied Psychology Unit, Department of Psychology, The University, Sheffield S10 2TN.

For *The British Journal of Mathematical and Statistical Psychology* should be sent to Dr. C. Leach, Department of Psychology, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

For *The British Journal of Educational Psychology* should be sent to Professor Dennis Child, School of Education, St. Thomas' Street, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU.

Contributors are advised to consult *Suggestions to Contributors* issued by the Standing Committee on Publications of the British Psychological Society and obtainable, price £2.00 post free, from The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

All manuscripts should be submitted in duplicate.

OFFPRINTS

Contributors on application receive fifty copies of their papers free. Additional copies may be obtained in units of fifty copies.

SUBSCRIPTIONS TO THE BRITISH PSYCHOLOGICAL SOCIETY

The subscription for Fellows, Associates and Graduate Members is £20.00.

They are entitled to purchase any or all of the Society's Journals at the special rates of £4.00 per Journal per annum.

Information on membership and activities may be obtained from the Secretary, The British Psychological Society, St. Andrew's House, 48 Princess Road East, Leicester, LE1 7DR.

BRITISH JOURNAL OF EDUCATIONAL PSYCHOLOGY,
VOL. 50, PART 3, 1980

CONTENTS

ARTICLES:

HEWISON, J. and TIZARD, J.	Parental Involvement and Reading Attainment	120
EVANS, L.	WISC Performance Scale and Coloured Progressive Matrices with Deaf Children	2
MCDERMOTT, P. A.	Principal Components Analysis of the Revised Bristol Social Adjustment Guides	2
FONTANA, D. and EVANS, H.	Mode of Stimulus Presentation and Short-term Memory Efficiency in Primary School Children	229
HORNSBY, B. and MILES, T. R.	The Effects of a Dyslexia-centred Teaching Programme	236
ROGERS, C. G.	The Development of Sex Differences in Evaluations of Others' Successes and Failures	243
CATTELL, R. B.	The Heritability of Fluid (g_f) and Crystallised (g_c) Intelli- gence Estimated by a Least Squares Use of the MAVA Method	25
WEBB, N. M.	An Analysis of Group Interaction and Mathematical Errors in Heterogeneous Ability Groups	266
RUSSELL, G. W. and LAMBERT, W. B.	Sherlock Holmes Meets Othello: A MDS Analysis of Literary Characters	277
MAQSUD, M.	Effects of Personal Lecture Notes and Teacher-notes on Recall of University Students	289

RESEARCH NOTES:

IVES, S. W.	The Use of Orientations in Children's Drawings of Familiar Objects: Principles versus Percepts	295
FRANCIS, L. and CARTER, M.	Church Aided Secondary Schools, Religious Education as an examination subject and Pupil Attitude towards Religion	297
NISBET, J., ADAMS, M. and ARTHUR, J.	A simple Form of Tailored Testing	301
HARTLEY, J., KENELY, J., OWEN, G. and TRUE- MAN, M.	The Effect of Headings on Children's Recall from Prose Text	304

BOOK REVIEWS	308
PUBLICATIONS RECEIVED	320
AUTHOR INDEX TO VOLUME 50	323
NOTES TO CONTRIBUTORS	325



UK ISSN 0007-0998

9 MAR 1981